目　录

第一部分　线性代数

第二部分　对称变换群

第三部分　矩阵论及其应用

第 一 部 分

线 性 代 数

第一章 向量空间

1.1 引论

我们通过叙述向量空间的概念开始代数的学习. 读者可能因为在物理学中碰到过向量而熟悉它们了. 它们被大小与方向所确定,因而从本质上说是几何量. 但是,如果我们希望解决向量方面的一些问题, 仅在纸上画箭头是不够的. 我们必须学会按某些规则进行向量的运算, 而这些规则又必须与物理学的经验一致. 我们将介绍一种向量的代数, 它仅仅是建立在向量加法与纯量乘法这两种运算的基础上的. 这将导致我们把向量空间定义为一个数学系统, 其代数性质类似于物理学中的向量.

数学中向量空间的例子俯拾皆是, 其中不少还是科学工作者感兴趣的. 我们在这一章及后面几章中将看到这一点.

1.2 向量

由大小和方向所刻画的物理量,例如力、位移、速度和加速度,称为向量. 在图上,向量用一个带箭头的线段表示,箭头确定了它的方向,而根据给定的尺度,用线段的长度表示向量的大小. 以下我们将用黑体字表示向量, 例如 $\mathbf{v}$. 若 $\mathbf{v}$ 的起点与终点分别为 A 与 B,则我们还可以记

$$\mathbf{v}=\overrightarrow{AB}.$$

两个长度相等、方向相同的带箭头的线段确定同一个向量,而不管它们在纸上的位置怎样. 于是,在图 1 中

$$\overrightarrow{AB}=\mathbf{v}=\mathbf{u}.$$

虽然向量本质上是几何量，但在这里我们所关心的却是向量的代数性质，亦即，在我们即将引入的某些合成规则（向量加法与减法、纯量乘法）下，向量的有关性质.

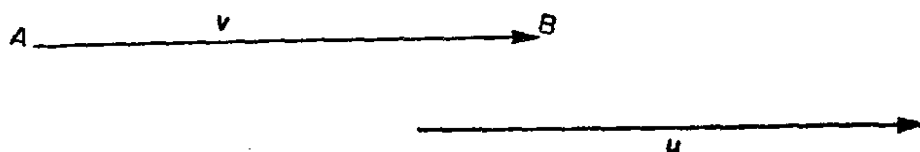


图 1　向量的相等

1.3　向量代数的法则

法则 I

按照向量加法的三角形法则可以把两个向量 **v** 及 **w** 相加，得出唯一的第三向量 **v**+**w**，它称为 **v** 与 **w** 的和(图 2)．把 **w** 的起点放到 **v** 的终点上，则从 **v** 的起点画到 **w** 终点的箭头表示 **v**+**w**．即使 **v** 和 **w** 是同向（或反向）的向量，这个法则仍适用，虽然在这种情况我们得到的三角形是退化的.



图 2　向量加法

在"**v**+**w**"中的加号表示向量加法，而且这个加号和用于加数的加号有相当不同的功用．严格地说，我们应当发明一个新符号来表示向量加法．然而，在不致产生混淆的情况下，特别是为了强调这些运算的共同特点时，数学家们乐意用同一个符号来表示不同的运算．在法则 II 和 III 中给出了两种加法的重要的共同性

质.

法则 II

向量加法是可交换的,即对一切向量 $\mathbf{v},\mathbf{w}$ 有

$$\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v}.$$

由图 3，这是显然的.

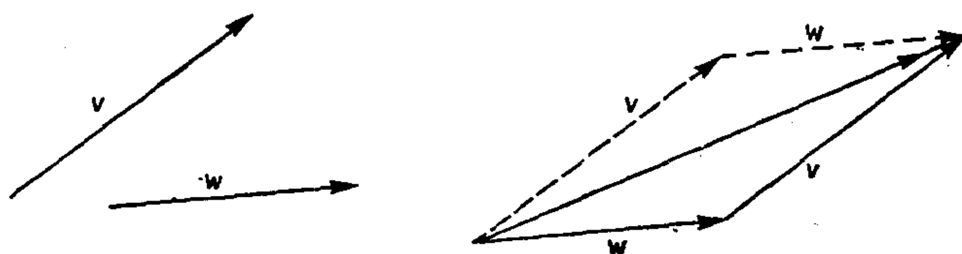


图 3 交换率: $\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v}$

法则 III

向量加法是可结合的,即对任意三个向量 $\mathbf{u},\mathbf{v},\mathbf{w}$

$$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w}).$$

这里括号指明了进行加法的次序. 为了证明结合律，参看图 4.

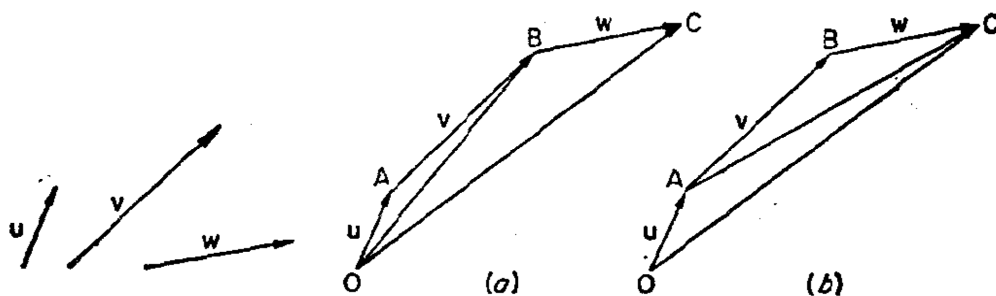


图 4 结合律: $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$

这个法则表明,当求三个向量 $\mathbf{u},\mathbf{v},\mathbf{w}$ 的和的时候，先求 $\mathbf{u}$ 与 $\mathbf{v}$ 的和还是先求 $\mathbf{v}$ 与 $\mathbf{w}$ 的和都可以. 于是我们可以把和写作 $\mathbf{u}+\mathbf{v}+\mathbf{w}$ 而不会有异义.

依次用 $\overrightarrow{OA}$, $\overrightarrow{AB}$ 与 $\overrightarrow{BC}$ 代表向量 $\mathbf{u}$, $\mathbf{v}$ 与 $\mathbf{w}$. 图 4(a) 说明

了 $\mathbf{w}$ 加$(\mathbf{u}+\mathbf{v})$(用 $\overrightarrow{OB}$ 代表)的加法，而在图 4(b)中，是把$(\mathbf{v}+\mathbf{w})$(用 $\overrightarrow{AC}$ 代表)加到 $\mathbf{u}$ 上．两种情形下结果都是向量 $\overrightarrow{OC}$．

设 $\mathbf{v}$ 和 $\mathbf{v}'$ 是有相同长度而方向相反的向量（图 5）．由三角形法则，$\overrightarrow{OC}$ 表示 $\mathbf{v}+\mathbf{v}'$，而 C 和 O 是重合的．于是 $\mathbf{v}+\mathbf{v}'$ 是长度为零的向量(并因此没有特定的方向)．下面的法则给出了这个向量的代数性质．

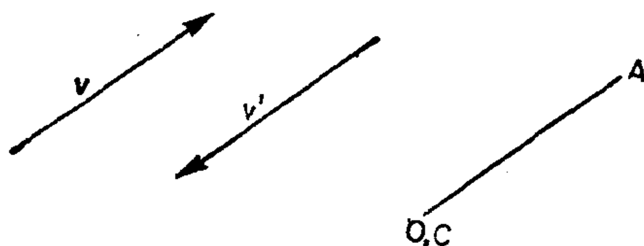


图 5

法则 IV

存在一个零向量 $\mathbf{0}$，对一切 $\mathbf{v}$ 有性质

$$\mathbf{v}+\mathbf{0}=\mathbf{0}+\mathbf{v}=\mathbf{v}.$$

图 5 的向量 $\mathbf{v}$，$\mathbf{v}'$ 满足代数法则

$$\mathbf{v}+\mathbf{v}'=\mathbf{0}.$$

和普通算术相似，自然地可记

$$\mathbf{v}'=-\mathbf{v}.$$

法则 V

对每个向量 $\mathbf{v}$ 存在一个向量 $-\mathbf{v}$，满足

$$\mathbf{v}+(-\mathbf{v})=(-\mathbf{v})+\mathbf{v}=\mathbf{0}.$$

$-\mathbf{v}$ 称为 $\mathbf{v}$ 关于加法的逆向量，或者 $\mathbf{v}$ 的负向量．

现在我们可以用

$$\mathbf{v}-\mathbf{w}=\mathbf{v}+(-\mathbf{w})$$

来定义向量减法．图 6 说明了这一点．

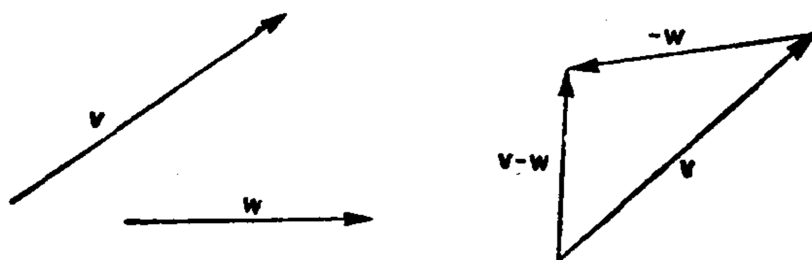


图 6 向量减法

向量减法很象普通减法．例如，如果我们已知

$$\mathbf{u}+\mathbf{v}=\mathbf{w},$$

我们可由这些法则断言

$$\mathbf{v}=\mathbf{w}-\mathbf{u}.$$

因为

$$\begin{aligned}
\mathbf{w}-\mathbf{u}&=\mathbf{w}+(-\mathbf{u}) && \text{(根据定义)}\\
&=(\mathbf{u}+\mathbf{v})+(-\mathbf{u}) && \\
&=\mathbf{u}+(\mathbf{v}+(-\mathbf{u})) && \text{(根据结合律)}\\
&=\mathbf{u}+(-\mathbf{u}+\mathbf{v}) && \text{(根据交换律)}\\
&=(\mathbf{u}+(-\mathbf{u}))+\mathbf{v} && \text{(又根据结合律)}\\
&=\mathbf{0}+\mathbf{v} && \text{(根据法则 V)}\\
&=\mathbf{v} && \text{(根据法则 IV)}
\end{aligned}$$

法则 VI

可以用一个数(纯量)k 来乘一个向量 $\mathbf{v}$，得到一个向量 $k\mathbf{v}$.

向量 $k\mathbf{v}$ 的大小为 $\mathbf{v}$ 的大小的 $|k|$ 倍．当 k 为正的时，$k\mathbf{v}$ 与 $\mathbf{v}$ 同向；当 k 为负的时，$k\mathbf{v}$ 与 $\mathbf{v}$ 反向(图 7)．

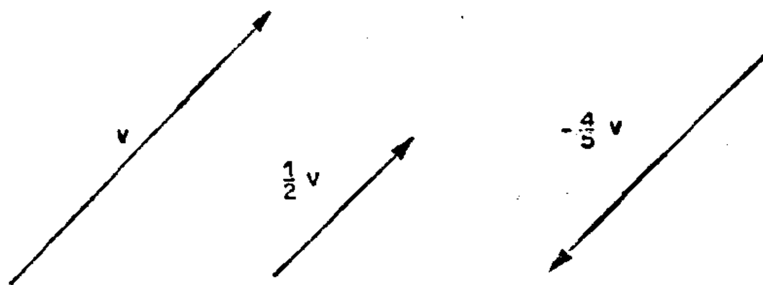


图 7　向量的纯量乘法

法则 VII

下面的运算规则对任何纯量和向量成立：

(i) $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$　　对向量加法的分配律．

(ii) $(k+l)\mathbf{v}=k\mathbf{v}+l\mathbf{v}$　　对纯量加法的分配律．

(iii) $(kl)\mathbf{v}=k(l\mathbf{v})$　　　　纯量乘法结合律.

(iv) $1\cdot\mathbf{v}=\mathbf{v}$　　　　纯量单位律.

用画图的办法容易看出向量确实满足这些运算规则.

向量代数法则一览

现在我们来总结向量代数的法则，不再引用在物理向量 相应情形中它们的几何意义. 这种纯粹代数抽象是向量空间一般定义的基础.

法则 I 到 V 关于向量加法

I. 可以把两个 向 量 $\mathbf{v}$ 与 $\mathbf{w}$ 相加，得到 称为和的第三个向量 $\mathbf{v}+\mathbf{w}$. 这称为向量加法的封闭律.

II. 向量 加法是可交换的.

III. 向量 加法是可结合的.

IV. 存在一个零向量，用 $\mathbf{0}$ 表示，对一切 $\mathbf{v}$ 有性质

$$\mathbf{v}+\mathbf{0}=\mathbf{0}+\mathbf{v}=\mathbf{v}.$$

V. 对任何向量 $\mathbf{v}$，存在一个相应的向量 $-\mathbf{v}$，称为 $\mathbf{v}$ 的负向量或逆向量，有性质

$$\mathbf{v}+(-\mathbf{v})=-\mathbf{v}+\mathbf{v}=\mathbf{0}.$$

法则 VI 和 VII 关于纯量乘法

VI. 一个纯量 k 和一个向量 $\mathbf{v}$ 的积 $k\mathbf{v}$ 本身是一个向量. 这称为纯量乘法的封闭律.

VII. 纯量乘法满足分配律

(i) $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$

和

(ii) $(k+l)\mathbf{v}=k\mathbf{v}+l\mathbf{v}$,

满足结合律

(iii) $(kl)\mathbf{v}=k(l\mathbf{v})$,

以及纯量单位律

(iv) $1\cdot\mathbf{v}=\mathbf{v}$.

上述 7 条法则好象是从向量服从的规则里相当随意挑选出来的．其实我们的挑选有两条理由．首先，这些法则足以完全描述向量在向量加法与纯量乘法下所具有的性质，任何这方面的性质都是这七个法则的推论．其次，这些法则容易用群这一概念的语言来记忆．我们后面将有一章研究群的概念．向量代数的头五条法则简单地叙述为在向量加法下向量的集合构成一个交换群．于是，用群的概念，我们可以把头五条法则合为一条，我们的向量代数法则就缩减为三条了．

现在我们把向量空间定义为某种量的一个集合 $\mathscr{V}$，在此集合中，可以把两个量相加或用纯量乘一个量，并满足法则 I 到 VII．$\mathscr{V}$ 的元素称为 $\mathscr{V}$ 的向量．

例 1 一元函数

设 $\mathscr{V}$ 是在一切数 x 上都有定义的所有函数的集合．

例如，由

$$f(x)=\sin x$$

与

$$g(x)=\begin{cases}x^2 (x\geqslant 0)\\ 1\ \ (x<0)\end{cases}$$

定义的函数 f 和 g 属于 $\mathscr{V}$．我们可以定义 f 和 g 的和 $f+g$，它在 x 上的值是

$$(f+g)(x)=f(x)+g(x).$$

我们还可以用一个数(纯量)k 乘 f 而得到函数 kf，定义

$$(kf)(x)=kf(x).$$

如果把 $\mathscr{V}$ 中的函数当作向量，那末容易验证 $\mathscr{V}$ 是适合法则 I 到 VII 的．(定义零函数 θ 为：$\theta(x)=0$，对一切 x；定义 f 的负函数为 $-f=(-1)f$.)于是 $\mathscr{V}$ 是一个向量空间．

如果向量空间 $\mathscr{V}$ 中的向量的一个集合 $\mathscr{S}$ 本身也构成一个向量空间（对于相同的向量加法和纯量乘法），那末 $\mathscr{S}$ 称为 $\mathscr{V}$

的子空间，且记作 $\mathscr{S}\subseteq\mathscr{V}$．如果 $\mathscr{S}$ 不包含 $\mathscr{V}$ 的每个向量，那末称 $\mathscr{S}$ 为 $\mathscr{V}$ 的一个真子空间，且记作 $\mathscr{S}\subset\mathscr{V}$．

例 2 连续函数

设 $\mathscr{V}$ 的定义如例 1．$\mathscr{S}$ 是 $\mathscr{V}$ 中一切连续函数的集合．于是例 1 的 f 属于 $\mathscr{S}$ 而 g 不属于 $\mathscr{S}$．现在我们来证明 $\mathscr{S}$ 是 $\mathscr{V}$ 的真子空间．法则 I 在 $\mathscr{S}$ 里成立，因为两个连续函数的和也是连续函数，因此属于 $\mathscr{S}$．法则 V 和 VI 因相似的理由而成立．零函数 θ 属于 $\mathscr{S}$，因此法则 IV 成立．剩下的法则是关于运算方法的，而因为对 $\mathscr{V}$ 它们是成立的，所以对 $\mathscr{S}$ 不必再验证了．最后，$\mathscr{V}$ 中存在不属于 $\mathscr{S}$ 的函数．于是 $\mathscr{S}\subset\mathscr{V}$．

例 3 微分方程的解集合

简谐运动的微分方程为

$$\frac{d^2y}{dt^2}+\omega^2y=0．(\omega\neq0),$$

它的所有解的集合 $\mathscr{V}$ 构成一个向量空间．假定 y_1 和 y_2 属于 $\mathscr{V}$，则

$$\frac{d^2}{dt^2}(y_1+y_2)+\omega^2(y_1+y_2)=\left(\frac{d^2y_1}{dt^2}+\omega^2y_1\right)+\left(\frac{d^2y_2}{dt^2}+\omega^2y_2\right)=0,$$

因此 y_1+y_2 属于 $\mathscr{V}$．这样，法则 I 成立．类似地可证明法则 IV，V，VI 成立．（$\mathscr{V}$ 里零向量是零函数 θ．）这就得出 $\mathscr{V}$ 是一个向量空间．我们没有解这微分方程就证得了这点．大家知道这方程的通解是

$$y(t)=a\sin\omega t+b\cos\omega t,$$

其中 a 和 b 是任意常数．

例 4 物理中的偏微分方程

物理中出现的三个最重要偏微分方程是：

(i) 拉普拉斯方程

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0.$$

例如，在一个不带电的空间区域中，静电位 $V(x, y, z)$ 就满足这个方程.

(ii) 波动方程

$$\nabla^2 U - \frac{1}{v^2} \frac{\partial^2 U}{\partial t^2} = 0.$$

描述波的运动的函数 $U(x, y, z, t)$ 适合此方程.

(iii) 薛定谔方程

$$\frac{-h^2}{2m} \nabla^2 \Psi + (V - E) \Psi = 0.$$

它是量子理论的基础. 这里 $\Psi(x, y, z)$ 描述了具有质量 m 与能量 E 的电子在电位 $V(x, y, z)$ 影响下的运动情况.

在以上三个微分方程中，每一个方程所有解的集合都构成一个向量空间.

练　　习

1. 证明微分方程

$$\frac{d^2 y}{dt^2} + \omega^2 y = 1$$

的解集合不构成一个向量空间.

[提示：证明违反了七条法则中的某一条即可，不必解微分方程.]

2. 由法则证明

$$-(\mathbf{v} - \mathbf{w}) = \mathbf{w} - \mathbf{v},$$

并用图说明.

3. 证明向量减法不可交换也不能结合.

4. 证明一个向量空间中只有一个零向量.

[提示：假定 $\mathbf{0}$ 和 $\mathbf{0}'$ 是两个零向量. 用两种方法计算 $\mathbf{0} + \mathbf{0}'$.]

5. 证明对向量空间的任意一个向量，只存在唯一的一个负向量.

[提示：假定 $\mathbf{v}'$ 和 $\mathbf{v}''$ 都是 $\mathbf{v}$ 的负向量. 用两种方法计算 $\mathbf{v}' + \mathbf{v} + \mathbf{v}''$.]

6. 由法则证明 $k\mathbf{0}=\mathbf{0}$ 及 $0\mathbf{v}=\mathbf{0}$.

[提示：要证明第一个等式，只要将分配律用于关系式 $k\mathbf{v}=k(\mathbf{v}+\mathbf{0})$.]

7. 证明若 $k\mathbf{v}=\mathbf{0}$,则 $k=0$ 或 $\mathbf{v}=\mathbf{0}$.

8. 用法则验证

$$3\mathbf{v}=\mathbf{v}+\mathbf{v}+\mathbf{v}.$$

请注意，向量 $\frac{1}{2}\mathbf{v}$ 没有这种解释.

9. 证明

$$-\mathbf{v}=(-1)\mathbf{v},$$

即用纯量-1 乘 $\mathbf{v}$ 得到 $\mathbf{v}$ 的负向量.

[提示：展开$(-1+1)\mathbf{v}$.]

10. 设 $\mathscr{V}$ 是向量空间,而 $\mathscr{S}$ 是 $\mathscr{V}$ 的向量的集合.证明如果 $\mathscr{S}$ 的向量对向量加法和纯量乘法满足封闭律，则 $\mathscr{S}$ 是 $\mathscr{V}$ 的一个子空间.

11. 设 $\mathscr{V}$ 是画在一个平面中的所有向量的向量空间. 单位长度的向量子集 $\mathscr{S}$ 构成 $\mathscr{V}$ 的一个子空间吗？

12. 找出尽可能多的向量空间例子. 回忆向量空间两个本质特点是，定义了(i) 向量加法,(ii) 纯量乘法,且使得在这些运算下,该系统是封闭的(法则 I 和 VI). 剩下的法则通常是容易检验的.

1.4 用向量方法解某些问题

用向量方法解力学问题

例 5 B 市在 A 市北偏东 30° 方向距 A 400 英里处. 一天，东风以 40 英里/小时稳定地刮着，一架飞机用最大速度花了 2 小时从A 飞到 B. 如果风向、风速保持不变,飞机往回飞行可能的最短时间是多少？

在 A 到 B 的飞行中，用大小为 200 英里/小时（两小时飞了400英里),方向为北偏东 30° 的向量 $\mathbf{u}$ 给出合速度. 我们有

$$\mathbf{u}=\mathbf{w}+\mathbf{v},$$

这里 $\mathbf{v}$ 表示无风时的飞机速度,而 $\mathbf{w}$ 表示风的速度(图 8).

我们用 $\|\mathbf{r}\|$ 表示向量 $\mathbf{r}$ 的大小. 为了简化计算,让我们以 40

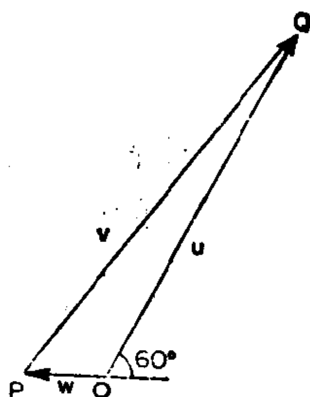


图 8 向外飞行

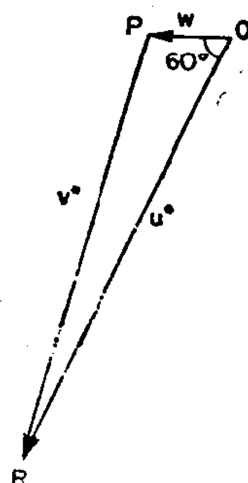


图 9 回程飞行

英里/小时为单位来度量，这样 $\|\mathbf{w}\|=1$ 且 $\|\mathbf{u}\|=5$. 于是由余弦定理

$$\begin{aligned}\|\mathbf{v}\|^2&=\|\mathbf{u}\|^2+\|\mathbf{w}\|^2-2\|\mathbf{u}\|\|\mathbf{w}\|\cos 120^\circ\\&=25+1+5=31.\end{aligned}$$

因此

$$\|\mathbf{v}\|=\sqrt{31}.$$

现在考虑回程飞行(图 9). 这里由

$$\mathbf{u}^*=\mathbf{w}+\mathbf{v}^*$$

给出飞机的合速度，其中 $\mathbf{v}^*$表示静止空气中飞机速度. 因为回程飞行还要求尽可能地快，所以我们有

$$\|\mathbf{v}^*\|=\|\mathbf{v}\|=\sqrt{31}.$$

(然而要注意，$\mathbf{v}$ 和 $\mathbf{v}^*$不是相等的向量，因为它们有不同的方向.)

由余弦定理

$$\|\mathbf{v}^*\|^2=\|\mathbf{u}^*\|^2+\|\mathbf{w}\|^2-2\|\mathbf{u}^*\|\|\mathbf{w}\|\cos 60^\circ,$$

因此

$$31=x^2+1-x,$$

这里 $x=\|\mathbf{u}^*\|$. 于是

$$0=x^2-x-30=(x-6)(x+5).$$

正值解给出了$\|\mathbf{u}^*\|$的值．故$\mathbf{u}^*$的大小为6×40英里/小时，而回程飞行花了$\frac{400}{240}$小时，即1小时40分钟．

用于解上述问题的方法是完全刻板的，读者有理由怀疑，解这样简单一个问题是否需要如此繁复的计算．的确可能有一个较易的解法（见本节末的练习13）．

用向量方法解几何问题

几何中许多问题都可归结为向量问题．我们在以O为原点的平面中，可以用向量$\overrightarrow{OA}$来表示一点A，这个向量称为点A对于原点O的位置向量．

例6 求用比$r:s$分线段AB的点P的位置向量公式(图10)．

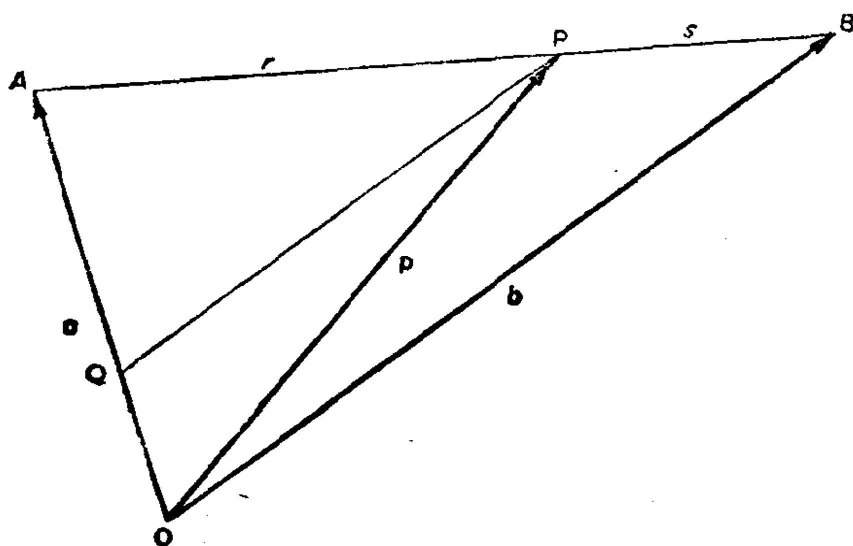


图10 线段AB的内分点

我们已知$AP:PB=r:s$．设$\mathbf{a},\mathbf{b}$依次是A,B的位置向量．在OA上确定Q使得QP平行于OB，则

$$QP=\frac{r}{r+s}OB,$$

且

$$OQ=\frac{s}{r+s}OA.$$

因此运用 $\mathbf{p}=\overrightarrow{OP}=\overrightarrow{OQ}+\overrightarrow{QP}$，我们得到公式

$$\mathbf{p}=\frac{s}{r+s}\mathbf{a}+\frac{r}{r+s}\mathbf{b}.$$

把这公式用于 $r=s=1$ 情形，我们求得 AB 中点的位置向量是 $\frac{1}{2}(\mathbf{a}+\mathbf{b})$.

例 7 证明三角形中线相交于一点，且交点是每条中线的三等分点.

在三角形 ABC 中(图 11)，设 M 是 CB 中点. 设 $\mathbf{a},\mathbf{b},\mathbf{c}$ 依次是 A,B,C 的位置向量(对于一个固定原点 O)，则 M 的位置向量是

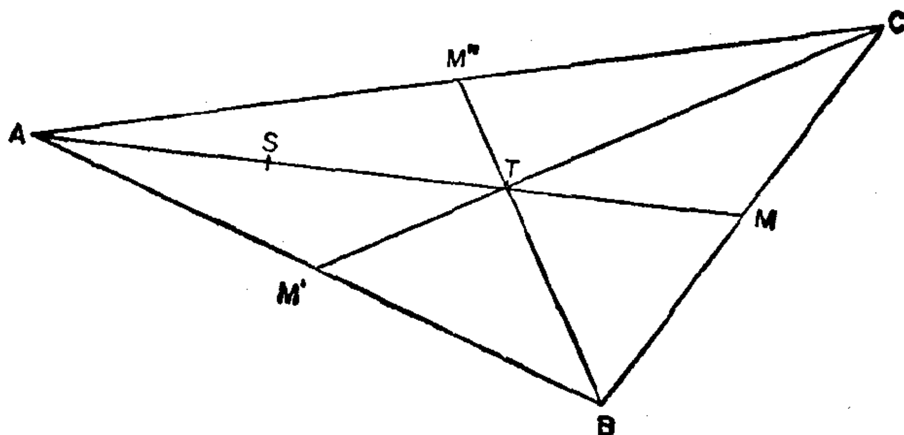


图 11 三角形中线的共点

$$\mathbf{m}=\frac{1}{2}(\mathbf{b}+\mathbf{c}),$$

同时 AM 的三等分点 S 和 T 依次有位置向量

$$\mathbf{s}=\frac{2}{3}\mathbf{a}+\frac{1}{3}\mathbf{m}$$

和

$$\mathbf{t}=\frac{1}{3}\mathbf{a}+\frac{2}{3}\mathbf{m}.$$

在后一式中把 $\mathbf{m}$ 代换掉，我们得到

$$\mathbf{t}=\frac{1}{3}\left(\mathbf{a}+\mathbf{b}+\mathbf{c}\right),$$

这个式子关于 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是对称的. 于是，由

$$\mathbf{t}'=\frac{1}{3}\left(\mathbf{a}+\mathbf{b}+\mathbf{c}\right)=\mathbf{t}$$

给出了中线 CM' 靠近底 AB 的三等分点 T' 的位置向量 $\mathbf{t}'$. 类似地 T 是从 B 所画中线的三等分点. 这证明中线共点，问题就解决了.

上述问题里选择原点 O 是非常随意的. 以下，我们将从图中省略掉原点，而且只说点的位置向量而不再说所参照的原点.

上面求出的三角形中线的交点 T 可以看成(i)均匀三角形薄板 ABC 的质量中心，以及(ii)由在 A,B,C 位置上的三个等质量质点组成的质点系的质量中心.

简单的质量中心问题常常最好是用向量方法解决. 在下例中我们要记住，如果把质量为 m 与 M 的两个质点分别放在 A,B 两点，则此质点系的质量中心 G 以 $M:m$ 之比分割线段 AB(图 12).

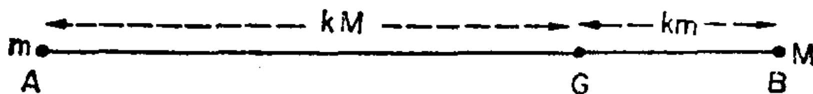


图 12　两个质点的质量中心

利用 A，B 的位置向量，G 的位置向量是

$$\mathbf{g}=\frac{m}{m+M}\mathbf{a}+\frac{M}{m+M}\mathbf{b}.$$

质量中心

例 8　设一个质点系由位于 A,B,C,D 的四个质点组成，它们的质量分别为 2,1,3,2 个单位. 求此质点系的质量中心(图

13).

如果P点为在A，B的质点组成的质点系的质量中心，Q点为在C，D的质点组成的质点系的质量中心，则原质点系等价于在P点放3个单位质量而在Q点放5个单位质量所成的质点系．相应的位置向量的关系是

$$\mathbf{p}=\frac{2}{3}\mathbf{a}+\frac{1}{3}\mathbf{b},$$

$$\mathbf{q}=\frac{3}{5}\mathbf{c}+\frac{2}{3}\mathbf{d}.$$

如果G点为原质点系的质量中心，则由在P、Q的两质点组成的质点系又可等价于在G点放8个单位质量．因此，G的位置向量是

$$\begin{aligned}\mathbf{g}&=\frac{3}{8}\mathbf{p}+\frac{5}{8}\mathbf{q}\\&=\frac{1}{8}\left(2\mathbf{a}+\mathbf{b}+3\mathbf{c}+2\mathbf{d}\right).\end{aligned}$$

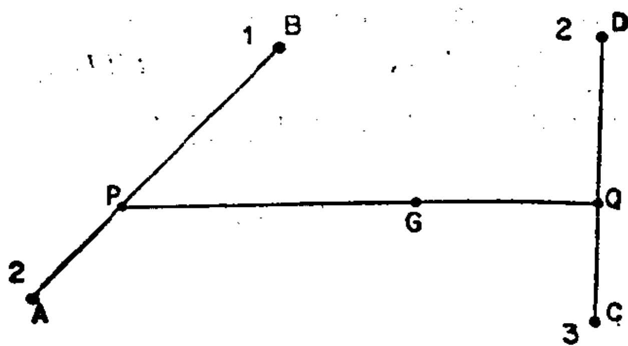


图 13　四质点质量中心

不难证明以下的推广：位于$A_1, A_2, \cdots, A_r$质量依次为$m_1, m_2, \cdots, m_r$的质点系统的质量中心是这样一点G，对于任意一个原点，相应位置向量的关系式

$$\mathbf{g}=\frac{1}{m_1+m_2+\cdots+m_r}\left(m_1\mathbf{a}_1+m_2\mathbf{a}_2+\cdots+m_r\mathbf{a}_r\right)$$

给出了G。

例 9 求一均匀固体的质量中心，这固体的形状是以A，B，C，D为顶点的四面体(图 14)。

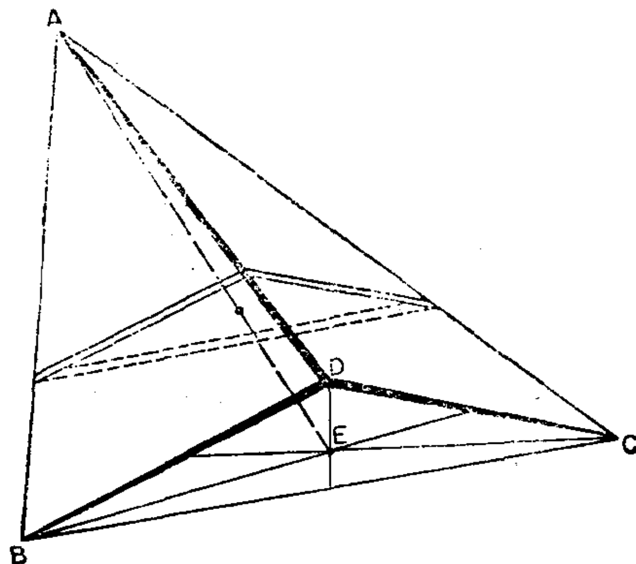


图 14 四面体质量中心

把这四面体切为平行于“底”BCD的三角形薄片并用在薄片质量中心的等价质点来代替每个薄片. 所有这些质点在连线AE上，E是底部薄片的质量中心，即三角形BCD的中线交点. 因此四面体质量中心G也一定在AE上. 注意

$$\mathbf{e}=\frac{1}{3}\left(\mathbf{b}+\mathbf{c}+\mathbf{d}\right)$$

给出E的位置向量.

因此G的位置向量必定形如

$$\mathbf{g}=\frac{r}{r+s}\mathbf{a}+\frac{s}{(r+s)}\,\frac{1}{3}\left(\mathbf{b}+\mathbf{c}+\mathbf{d}\right),$$

或更简单地，对某常数$k(0\leqslant k\leqslant 1)$

$$\mathbf{g}=k\,\mathbf{a}+\frac{1}{3}\left(1-k\right)\left(\mathbf{b}+\mathbf{c}+\mathbf{d}\right). \qquad (1)$$

现在以顶点B代替顶点A，以“底”CDA代替BCD，根据类似的论证，我们断言G必在BF上，这里

$$\mathbf{f}=\frac{1}{3}\left(\mathbf{c}+\mathbf{d}+\mathbf{a}\right),$$

还断言，对某常数 l $(0\leqslant l\leqslant 1)$

$$\mathbf{g}=l\mathbf{b}+\frac{1}{3}(1-l)(\mathbf{c}+\mathbf{d}+\mathbf{a}).$$

用 $\mathbf{g}$ 的这两个表达式比较 $\mathbf{a}$ 和 $\mathbf{b}$ 的系数，我们求得 $k=l=\frac{1}{4}$，因此

$$\mathbf{g}=\frac{1}{4}(\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}). \tag{2}$$

这个表达式关于 $\mathbf{a},\mathbf{b},\mathbf{c}$ 和 $\mathbf{d}$ 是对称的．这是我们意料之中的，因为不存在位置向量 $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ 的任何一个系数不同于其余随便哪一个系数的理由．

我们可以把对称性用于表达式(1)，得出 $k=\frac{1}{3}(1-k)$，或 $k=\frac{1}{4}$，这和表达式(2)一致．

练　习

13．用下法解飞行问题(例 5)．把图 8 和图 9 拼成一个图（OP 是二者公共的）．可知因为 $\|\mathbf{v}\|=\|\mathbf{v}^*\|$，三角形 RPQ 是等腰的．证明

$$\|\mathbf{u}^*\|=\|\mathbf{u}\|+1 \text{ 单位}.$$

14．对一个以 12 英里/小时速度向南骑自行车的人来说，风好像从西边吹来．当他把他的速度增加到 21 英里/小时，风好像从西南吹来．求风的速度．

［提示：设人的速度是 $\mathbf{v}$，风的实际速度是 $\mathbf{w}$，则人所感觉到的风速是 $\mathbf{w}-\mathbf{v}$.］

15．$ABCDEF$ 是正六边形．求用向量 $\overrightarrow{AB}$，$\overrightarrow{AC}$，$\overrightarrow{AD}$，$\overrightarrow{AE}$ 和 $\overrightarrow{AF}$ 所表示力的合力．

［提示：把 $\overrightarrow{AB}$ 表示为 $\overrightarrow{AO}+\overrightarrow{OB}$，这里 O 是这六边形的中心，诸如此类．］

16. $ABCD$ 是平行四边形，而 M 是 BC 的中点. 用向量方法证明 AM 和 BD 互相三等分.

[提示：利用 A,B,C,D 的位置向量来表示 AM 和 BD 的三等分点的位置向量.]

17. $ABCD$ 是平行四边形，而 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是 A,B,C 的位置向量(对于某原点). 求 D 的位置向量.

18. 证明，联结四面体对边中点的三条线段共点且互相平分.

19. 依次置质量为 1,2,3,4 单位的质点于单位正方形 $ABCD$ 的四个顶点 A,B,C,D. 求这个系统的质量中心.

关于力学中向量方法进一步的练习可在 R. C. Smith & P. Smith, *Mechanics*, Wiley, 1968中找到.

1.5 在一个坐标系中的向量

考虑在同一平面上的一些向量(例如速度). 对于一个直角坐标系，当把向量 $\mathbf{v}$ 的起点放在原点 O，则可用 $\mathbf{v}$ 的终点坐标 (v_1, v_2) 表示向量 $\mathbf{v}$(图 15).

我们把向量和它的坐标表示等同起来且记作

$$\mathbf{v}=(v_1, v_2).$$

数 v_1, v_2 称为 $\mathbf{v}$ 关于给定坐标系的分量. 如果点 P 对于 O 的位置向量是 $\mathbf{v}$，那么 $\mathbf{v}$ 的分量就等于 P 点的坐标.

从向量加法的三角形法则我们得到

$$(v_1, v_2)+(w_1, w_2)=(v_1+w_1, v_2+w_2). \tag{3}$$

对任意纯量 k 还有

$$k(v_1, v_2)=(kv_1, kv_2). \tag{4}$$

图 16 所示的情形是

$$\mathbf{v}=(3,2), \quad \mathbf{w}=(-2,1), \quad \mathbf{v}+\mathbf{w}=(1,3), \quad k=-\frac{1}{2},$$

$$k\mathbf{v}=\left(-\frac{3}{2}, -1\right).$$

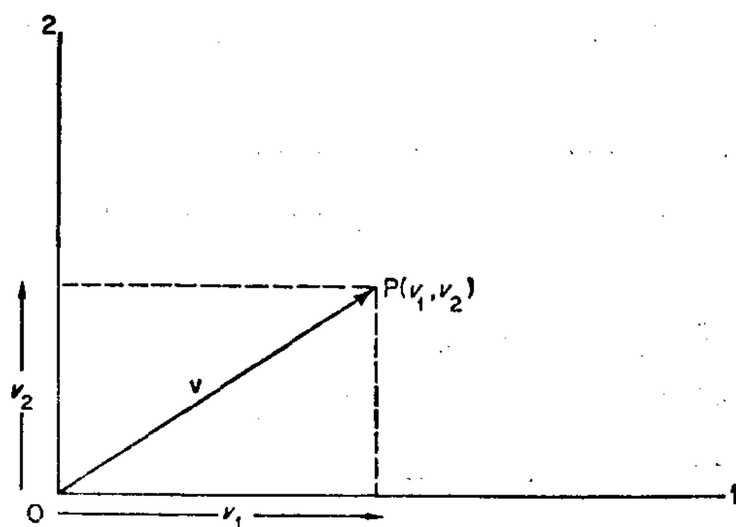


图 15 向量的分量

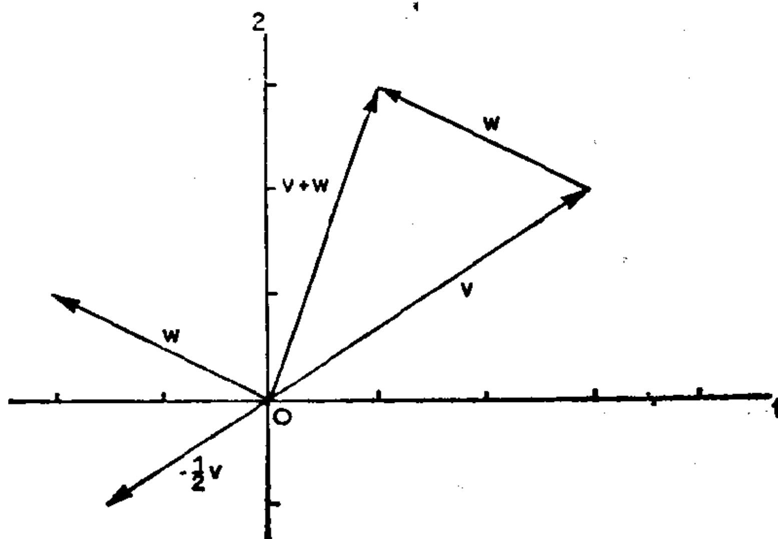


图 16

对于一切实数 v_1, v_2, 如果我们用(3)和(4)式定义坐标向量 (v_1,v_2) 的加法和纯量乘法，则 (v_1,v_2) 的系统构成一个向量空间．这个系统称实二-空间，并用 $\mathscr{R}^2$ 表示它（字母 $\mathscr{R}$ 代替"实分量"）．对于某些问题，允许采用复分量 v_1, v_2 和复纯量 k 是方便的（练习 28）．于是我们得到复二-空间 $\mathscr{C}^2$.

推而广之，则得重要的向量空间 $\mathscr{R}^n$ 和 $\mathscr{C}^n$，即实的和复的 n-空间．定义 $\mathscr{R}^n$($\mathscr{C}^n$) 为具有 n 个实（复）分量 $v_1, v_2, \cdots, v_n$ 的

一切 n 元组的系统，可按下列规则相加及用实（复）纯量乘这些 n 元组：

$$(v_1, v_2, \cdots, v_n) + (w_1, w_2, \cdots, w_n)$$
$$= (v_1 + w_1, v_2 + w_2, \cdots, v_n + w_n) \tag{5}$$

和

$$k(v_1, v_2, \cdots, v_n) = (kv_1, kv_2, \cdots, kv_n). \tag{6}$$

容易验证 $\mathscr{R}^n$ 和 $\mathscr{C}^n$ 是向量空间。零向量是 $(0, 0, \cdots, 0)$，$\mathbf{v} = (v_1, v_2, \cdots, v_n)$的负向量是

$$-\mathbf{v} = (-v_1, -v_2, \cdots, -v_n).$$

在 $\mathscr{R}^n$ 中纯量乘法采用实数，而在 $\mathscr{C}^n$ 中纯量乘法采用复数（它包含实数作为特殊情况）。因此，$\mathscr{R}^n$称为在实数上的向量空间，或者实向量空间，而 $\mathscr{C}^n$ 称为在复数上的向量空间，或复向量空间.

我们已经看到，对于给定的坐标系，可以把在一个平面里作用的物理向量表示为$\mathscr{R}^2$中的向量. 类似地，可以把在一个空间里作用的物理向量表示为 $\mathscr{R}^3$ 中的向量，对 $n>3$，我们画不出 $\mathscr{R}^n$ 中向量的草图. 尽管如此，在 $\mathscr{R}^n$ 及在 $\mathscr{C}^n$ 中引进长度和角度的一般概念仍然是可能的和有用的. 后面有一章中我们将要这样做.

练　　习

20. 设 $\mathbf{v} = (2, -1)$，$\mathbf{w} = (1, -1)$. 计算 $2\mathbf{v}$，$3\mathbf{w}$ 和 $2\mathbf{v} - 3\mathbf{w}$，并用草图说明。举出一个类似的例子来说明一个坐标系中的向量加法和纯量乘法.

21. 一个直角坐标系中，一个三角形顶点是$(0,2)$，$(-1,3)$和$(-2,-4)$，求这三角形中线交点坐标. 画草图.

22. 在一个直角坐标系中，已知一个单位正方形的顶点是 $A\ (0,0)$，$B\ (1,0)$，$C(1,1)$，$D(0,1)$，依次把质量为1,2,3,4单位的质点放在 A, B, C, D. 求质量中心坐标（和1.4节练习19比较一下）.

23. 一个均匀固体四面体的顶点是 $(1,-1,0)$，$(2,3,1)$，$(0,-1,1)$ 和 $(1,3,-2)$，求它的质量中心坐标.

24. 验证 $\mathscr{R}^n$ 满足向量代数七条法则.

25. 设 $\mathscr{R}^4$ 中 $\mathbf{v}=(2,-1,1,1)$，$\mathbf{w}=(-3,2,1,0)$。计算 $2\mathbf{v}$，$3\mathbf{w}$ 和 $2\mathbf{v}-3\mathbf{w}$。

26. 验证 $x_1=2, x_2=-2, x_3=1, x_4=1$ 是四元方程组

$$x_1+x_2+x_3-x_4=0$$
$$x_1+2x_2-x_3+3x_4=0$$

的解。可以方便地把这个解表示为 $\mathscr{R}^4$ 中的向量 $\mathbf{v}=(2,-2,1,1)$，其中第一个分量表示 x_1 取的值，第二分量表示 x_2 取的值，等等。验证 $\mathbf{w}=(-3,2,1,0)$ 是这方程组另一组解，$2\mathbf{v}$，$3\mathbf{w}$ 和 $2\mathbf{v}-3\mathbf{w}$ 也是这方程组的解。一般地证明，一切(实)解向量集合构成一个向量空间，它是 $\mathscr{R}^4$ 的真子空间。

[注意：不必求出这方程组的通解。]

27. 验证方程组

$$x_1+x_2+x_3-x_4=1$$
$$x_1+2x_2-x_3+3x_4=0$$

的解向量集合不构成向量空间。

[提示：参见法则 IV。]

28. 验证复系数方程组

$$(1+i)x_1+x_2=0$$
$$2x_1+(1-i)x_2=0$$

的解向量集合构成 $\mathscr{C}^2$ 的真子空间。

29. 下列 $\mathscr{R}^4$ 的向量集合中，哪些集合构成 $\mathscr{R}^4$ 的子空间？

(i) 所有 $v_2=0$ 的 (v_1,v_2,v_3,v_4)。

(ii) 所有 $v_2=1$ 的 (v_1,v_2,v_3,v_4)。

(iii) 所有适合 $v_3+2v_2=v_1$ 的 (v_1,v_2,v_3,v_4)。

第二章　线性方程组

2.1 引论

方程集合

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{array}\right\} \quad (1)$$

称为 n 元 m 个线性方程的方程组．方程个数 m 不一定等于未知数的个数 n．mn 个数 $a_{11}, a_{12}, \cdots, a_{mn}$ 称为方程组的系数，$b_1, b_2, \cdots, b_m$ 称为方程组的常数项，它们都假定是已知的．

方程组(1)的一个解是 n 个常数 $c_1, c_2, \cdots, c_n$ 的一个集合，使得我们在方程组(1)中用 c_1 代替 x_1，c_2 代替 x_2，$\cdots$，c_n 代替 x_n，每一个方程都满足．例如，三元两个方程的方程组

$$\begin{array}{l} 4x_1+7x_2+10x_3=-1 \\ 2x_1+3x_2+4x_3\ =-1 \end{array} \quad (2)$$

有解

$$x_1=-1,\quad x_2=-1,\quad x_3=1.$$

然而 $x_1=5, x_2=-3, x_3=0$ 不是这方程组的解，因为虽然这个三个值适合第一个方程，可是它们不适合第二个方程．

方程组一切解的集合称为这方程组的通解．方程组(2)的通解是

$$x_1=-2+k,\quad x_2=1-2k,\quad x_3=k,$$

这里 k 是任意的．取 $k=1$ 就得到上面所列举的解．它称为一个特解．

不是所有线性方程组都有解。方程组

$$\begin{aligned} x_1 + \ \ x_2 &= 1 \\ 2x_1 + 2x_2 &= 3 \end{aligned} \tag{3}$$

的两个方程显然是相互矛盾的，不管我们怎样取 x_1, x_2 的值. 方程组(3)称为不相容的. 至少有一个解的方程组称为相容的.

所有常数项 $b_1, \cdots, b_m$ 等于零的线性方程组称为齐次的. 齐次线性方程组总是相容的，因为它们总有平凡解 $x_i=0$，对一切 i. 可能有也可能没有其它解. 例如

$$\begin{aligned} x_1 + \ \ x_2 &= 0 \\ x_1 + 2\,x_2 &= 0 \end{aligned}$$

的唯一解是 $x_1 = x_2 = 0$. 方程组

$$\begin{aligned} x_1 + \ \ x_2 + x_3 &= 0 \\ x_1 + 2\,x_2 - x_3 &= 0 \end{aligned}$$

有非平凡解. 它的通解是

$$x_1 = -3\,k, \quad x_2 = 2\,k, \quad x_3 = k,$$

其中 k 任意.

现代计算机能方便地处理线性方程，即使所涉及方程组的方程个数和未知数个数是巨大的. 在这一章里，我们将给出一个解方程组的系统方法，即高斯化简法，方法是逐次消去未知数. 为了达到理论上的目的，这个方法是令人满意的，但在实际解方程组时，这个方法并不方便. 因此，实际使用的是改进了的方法.

例 1

让我们从第一个方程消去 x_1 来解方程组

$$\begin{aligned} 4\,x_1 + 7\,x_2 + 10\,x_3 &= -1 \\ 2\,x_1 + 3\,x_2 + \ \ 4\,x_3 &= -1. \end{aligned} \tag{2}$$

第一个方程减去第二个方程的两倍就办到了这一点. 于是，我们得到一个新方程组

$$x_2+2x_3=1$$
$$2x_1+3x_2+4x_3=-1. \tag{4}$$

现在,方程组(2)和(4)有相同的通解(见练习1和2.3节). 易见方程组(4) 的通解是

$$x_3=k\ (\text{任意}),\quad x_2=1-2k,$$
$$x_1=\frac{1}{2}(-1-4x_3-3x_2)=-2+k.$$

那末这也是 (2)的通解.

线性方程组的矩阵

阵列

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称为方程组(1)的矩阵. 它有m行和n列. 我们说A是一个$m\times n$矩阵. 数a_{ij}位于A的第i行和第j列上. 我们称a_{ij}为A中ij-元素.我们用(a_{ij}),这里$i=1,\cdots,m$和$j=1,\cdots,n$,作为一个矩阵的简略记号,当阵列中的行列数都明确时,我们就只用(a_{ij})来表示矩阵.

方程组(1)依赖于它的系数矩阵A和常数项$b_1,b_2,\cdots,b_m$.

阵列

$$A_b=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}$$

称为方程组(1)的增广矩阵. A_b是$m\times(n+1)$矩阵. 它完全确定了(1).

例 2

方程组

$$\begin{aligned} x_1+2x_2-3x_3&=-8\\ x_1+x_2+2x_3&=0 \end{aligned} \tag{5}$$

的增广矩阵是

$$\begin{pmatrix} 1 & 2 & -3 & -8 \\ 1 & 1 & 2 & 0 \end{pmatrix}.$$

它的行是(1 2 −3 −8)和(1 1 2 0). 它的列是

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix} \text{和} \begin{pmatrix} -8 \\ 0 \end{pmatrix}.$$

例 3

把表

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & 4 & -1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}$$

当作增广矩阵，便确定了线性方程组

$$\begin{aligned} 3x_1+x_2&=0\\ x_1+4x_2&=-1\\ 2x_1+x_2&=1\\ x_1+3x_2&=1. \end{aligned}$$

等价方程组

两个有相同通解的方程组称为等价的. 于是(2)和(4)是等价方程组. 它们的增广矩阵依次是

$$\begin{pmatrix} 4 & 7 & 10 & -1 \\ 2 & 3 & 4 & -1 \end{pmatrix}$$

和

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ 2 & 3 & 4 & -1 \end{pmatrix}.$$

如果我们把第一个矩阵的行当作 $\mathscr{R}^4$ 中的向量

$$\mathbf{v}_1=(4 \quad 7 \quad 10 \quad -1)$$

和

$$\mathbf{v}_2=(2 \quad 3 \quad 4 \quad -1),$$

则第二个矩阵的行是 $\mathbf{v}_1-2\mathbf{v}_2$ 和 $\mathbf{v}_2$. 这说明了利用增广矩阵的行描述方程的运算是方便的. 现在我们将引入三种(行)运算, 可用这三种运算把一个方程组化简为非常简单的等价方程组.

练　　习

1. 证明方程组(2)和(4)有相同的通解. 先证明方程组(2)的任意一个解也是方程组(4)的解.

[提示: 把(4)中第一个方程写作 $(4x_1+7x_2+10x_3)-2(2x_1+3x_2+4x_3)=-1-2(-1)$.]

类似地证明方程组(4)的任意一个解也是(2)的解. 注意不必为证明它们的等价性而解任何一个方程组.

2. 考察两个方程组

$$\begin{aligned} 4x_1+7x_2+10x_3&=-1 \\ 2x_1+3x_2+4x_3&=-1 \end{aligned} \tag{6}$$

和

$$4x_1+7x_2+10x_3=-1. \tag{7}$$

第二个方程组由单个方程组成. 显然方程组(6)的每个解也都是(7)的一个解. 证明反之不然.

[提示: 只要找出方程组(7)一个解, 它不适合方程组(6)就可以了. 方程组(6)和(7)不是等价的.]

3. 证明例2方程组(5)的通解是

$$x_3=k\ (\text{任意}),\quad x_1=8-7k,\quad x_2=-8+5k.$$

4. 证明如果齐次线性方程组有一个非平凡解, 则它有无穷多解.

5. 把基尔霍夫定律用于图17所示的电路, 对未知电流得到下列线性方

程组

$$
\begin{aligned}
I_1-I_2-I_3&=0\\
2I_2-6I_3&=0\\
I_1+2\quad I_2&=10.
\end{aligned}
$$

确定 I_1, I_2, I_3.

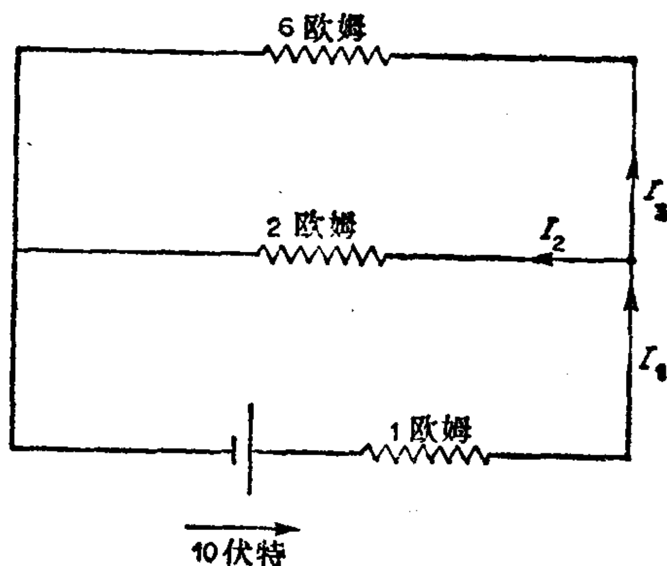


图 17 电路

2.2 行初等变换

作用在一个方程组上的以下三类变换称为初等变换:

类型 I:交换两个方程;

类型 II:用一个非零常数乘一个方程;

类型 III:把一个方程的 k 倍加到另一个方程上去.

在增广矩阵上的对应变换称为行初等变换. 只要把"方程"统统用"行"代替,上面的定义就变成了行初等变换的定义了. 注意当在一个方程组上实行一个类型 II 或类型III的初等变换时,只改变方程组中的一个方程,形成一个新方程组,它的方程除一个以外同原方程组方程相一致. 一个类型 I 的初等变换改变两个方程.

我们下一节要证明,通过(行)初等变换总是得到等价的方程

组．因此，如果我们用(行)初等变换来化简一个方程组，总能保证最终方程组的通解和原先方程组的通解是相同的．

我们举一个简单的例子来说明高斯化简法．我们注意到类型III的变换可用于消去未知数，其余两类变换可使我们得到一个特别简单的线性方程组，而这线性方程组的通解是显然的．

例 4 对增广矩阵进行行初等变换，把方程组(见2.1节(2))

$$4x_1+7x_2+10x_3=-1$$
$$2x_1+3x_2+4x_3=-1$$

化简，再写出这方程组的通解．

在下表中，我们并列给出方程组及其增广矩阵的代换步骤．(以后，将略去方程，因为显然，考察增广矩阵就足够了．)我们把矩阵的行加上标记，由这些标记应该可以清楚地看出所进行的行初等变换了．这些变换依次是：

表 1

矩阵	相应的方程组
$\begin{pmatrix}4&7&10&-1\\2&3&4&-1\end{pmatrix}\begin{matrix}a_1\\a_2\end{matrix}$	$\begin{matrix}4x_1+7x_2+10x_3=-1\\2x_1+3x_2+4x_3=-1\end{matrix}$
$\begin{pmatrix}0&1&2&1\\2&3&4&-1\end{pmatrix}\begin{matrix}b_1=a_1-2a_2\\b_2=a_2\end{matrix}$	$\begin{matrix}x_2+2x_3=1\\2x_1+3x_2+4x_3=-1\end{matrix}$
$\begin{pmatrix}0&1&2&1\\2&0&-2&-4\end{pmatrix}\begin{matrix}c_1=b_1\\c_2=b_2-3b_1\end{matrix}$	$\begin{matrix}x_2+2x_3=1\\2x_1-2x_3=-4\end{matrix}$
$\begin{pmatrix}0&1&2&1\\1&0&-1&-2\end{pmatrix}\begin{matrix}d_1=c_1\\d_2=\frac{1}{2}c_2\end{matrix}$	$\begin{matrix}x_2+2x_3=1\\x_1-x_3=-2\end{matrix}$
$\begin{pmatrix}1&0&-1&-2\\0&1&2&1\end{pmatrix}\begin{matrix}e_1=d_2\\e_2=d_1\end{matrix}$	$\begin{matrix}x_1-x_3=-2\\x_2+2x_3=1\end{matrix}$

(i) 把第二行的−2倍加到第一行上(类型III)，从第一个方程消去x_1．

(ii) 把第一行的-3倍加到第二行上(类型 III)，从第二个方程消去x_2。

(iii) 以$\frac{1}{2}$乘第二行(类型 II)。

(iv) 交换两行(类型 I)。

根据本题前的说明，原方程组的通解等同于最后方程组的通解，这个通解是

$$x_3=k\ (\text{任意}),\quad x_1=-2+k,\quad x_2=1-2k.$$

当然这通解从第三个矩阵起已经明显了. 要设计解线性方程组的一般过程，其困难之一是指出什么时候此过程应当停止. 我们希望能有一种以唯一的特殊形式的矩阵为其结果的解题方法. 上例中最后一个矩阵是所谓简化阶梯形的矩阵，它的定义如下.

定义 Ia 如果一个矩阵中每个非零行的首元素（第一个非零元素）出现在上一行首元素的右边，同时没有一非零行出现在零行之下，那末称这个矩阵是阶梯形的.

下列矩阵是阶梯矩阵的例子：

$$\begin{pmatrix} 1 & 6 & 2 & 4 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

和

$$\begin{pmatrix} 0 & 3 & -5 & 6 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

定义 Ib 一个矩阵称为是简化阶梯形的，如果

(i) 它是阶梯形；

(ii) 每一个非零行的首元素是 1；

(iii) 包含首元素列的其余元素都是 0.

例如以下两个矩阵，第一个是简化阶梯形的，第二个不是，因为不

满足条件(iii)：

$$\begin{pmatrix} 0 & 1 & -4 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

和

$$\begin{pmatrix} 0 & 1 & -4 & 2 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

显然用一系列行初等变换可以把任何矩阵化为一个简化阶梯矩阵．为此我们将在 2.4 节里给出一个系统的方法．此外，还能够证明，得到的简化阶梯矩阵是唯一的．于是这就为方程组增广矩阵化简过程提供了一个方便的目标．化简到这种形式以后，就可以毫无困难地写出方程组的通解了．

练　　习

6．用行初等变换把矩阵

$$\begin{pmatrix} 1 & 2 & -3 & -8 \\ 1 & 1 & 2 & 0 \end{pmatrix}$$

化为简化阶梯矩阵．由此得出 2.1 节方程组(5)的通解．

7．类似地解决电路问题(2.1 节练习 5)．注意只有唯一解．

8．证明方程组

$$\begin{aligned} 2x_1+3x_2-\ x_3+3x_4&=1 \\ 5x_1+\quad x_2\quad +7x_3&=-2 \\ -x_1+5x_2-9x_3+6x_4&=b \end{aligned}$$

当 $b=4$ 是相容的，而其它情形均不相容．

[提示：用初等变换从第一、二个方程消去x_1．]

2.3　行初等变换的理论

这一节里，我们将简略地研究一下行初等变换理论．特别，我

们将证明上一节所引用的性质，即通过（行）初等变换得到等价的方程组(见定理 I)。我们以后还将看到，行初等变换在矩阵论中占据重要的位置。并且除了方程论外它还有许多应用。

行初等变换的逆

先考察下面简单的例子。

例 5 在 4×2 矩阵 $A=(a_{ij})$ 上实行(类型 III 的)变换"把第二行的－5倍加到第三行上"。

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} \text{ 变为 } B=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31}-5a_{21} & a_{32}-5a_{22} \\ a_{41} & a_{42} \end{pmatrix}.$$

对每一个行初等变换 E 存在一个抵消它作用的另一个行初等变换 F：如果 E 把矩阵 A 变为 B，那末存在 F 把 B 变回到 A。F 称为 E 的逆。例如"把第二行的－5倍加到第三行上"的逆是"把第二行的 5 倍加到第三行上"。请在例 5 的矩阵 B 上验证这点。可把行初等变换和它的逆列成表2。

表 2

类型	行初等变换	逆变换
I	交换两行	交换同样的两行
II	用 $c\neq 0$ 乘一行	用 $1/c$ 乘同一行
III	把第 i 行的 k 倍加到第 j 行上	把第 i 行的 $-k$ 倍加到第 j 行上

行-等价

矩阵 A 行-等价于矩阵 B，记作 $A \underset{r}{\sim} B$，当且仅当用有限个行初等变换可把 A 化为 B。

我们已经看到行初等变换是可逆的。因此，如果 $A \underset{r}{\sim} B$ 则也

有 $B \underset{\sim}{r} A$. 故关系"$\underset{\sim}{r}$"是对称的. 它还是传递的：如果 $A \underset{\sim}{r} C$ 和 $C \underset{\sim}{r} B$ 则 $A \underset{\sim}{r} B$；又是自反的：对所有矩阵 A 有 $A \underset{\sim}{r} A$. 这三个性质使"$\underset{\sim}{r}$"成为一个等价关系. 这种关系在数学里是特别重要的. 作为一个简单的讨论，可见附录.

例 6 对矩阵

$$A=\begin{pmatrix} 4 & 3 & 0 \\ 2 & 1 & 1 \\ -1 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \text{和 } B=\begin{pmatrix} -2 & 36 & 20 \\ 2 & 1 & 1 \\ 3 & 3 & 4 \\ -1 & 3 & 2 \end{pmatrix},$$

用下列行初等变换把 A 化为 B 来证明 $A \underset{\sim}{r} B$：

(i) 把第三行的 5 倍加到第一行上；

(ii) 用 2 乘第一行；

(iii) 交换第三和第四行.

用这些步骤可得到如下一系列矩阵：

$$\begin{pmatrix} 4 & 3 & 0 \\ 2 & 1 & 1 \\ -1 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \underset{\sim}{r} \begin{pmatrix} -1 & 18 & 10 \\ 2 & 1 & 1 \\ -1 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \underset{\sim}{r} \begin{pmatrix} -2 & 36 & 20 \\ 2 & 1 & 1 \\ -1 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \underset{\sim}{r} \begin{pmatrix} -2 & 36 & 20 \\ 2 & 1 & 1 \\ 3 & 3 & 4 \\ -1 & 3 & 2 \end{pmatrix}.$$

线性方程组上的应用

定理 I 行-等价矩阵确定等价的方程组（即，有相同通解的方程组）.

这个定理是以下引理的简单推论.

引理 增广矩阵为

$$A_b=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

的线性方程组的每一个解，都是增广矩阵行-等价于 A_b 的方程组

的解.

证明 我们对 A_b 施行类型 III 的变换"把第 i 行的 k 倍加到第 j 行上"来论证引理. 原来的和新的方程组依次是

$$\left.\begin{array}{l}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{i1}x_1+\cdots\quad+a_{in}x_n=b_i\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{j1}x_1+\cdots\quad+a_{jn}x_n=b_j\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\end{array}\right\} \tag{8}$$

和

$$\left.\begin{array}{r}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{i1}x_1+\cdots\quad+a_{in}x_n=b_i\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_{j1}+ka_{i1})x_1+\cdots\quad+(a_{jn}+ka_{in})x_n=b_j+kb_i\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\end{array}\right\} \tag{9}$$

这两个方程组的不同仅仅在于第 j 个方程，新方程组的第 j 个方程是

$$(a_{j1}x_1+\cdots\quad+a_{jn}x_n)+k(a_{i1}x_1+\cdots\quad+a_{in}x_n)=b_j+kb_i.$$

显然方程组 (8) 的任何解都适合这个方程，因此也都适合方程组(9)的每个方程.

对类型 I 和类型 II 的变换，引理是显然的. 容易看出对一系列行初等变换引理是正确的.

现在可从引理和行-等价的对称性直接推出定理 I .

练　　习

9. 施行一系列行初等变换所得矩阵是依赖于所施行变换次序的. 把例8所列的行初等变换按如下顺序施行于A来说明上述断言： 先施行变换(ii)，再施行(iii)，然后是(i)，把所得结果和正文比较.

10. 设A和B是例6的矩阵. 证明$B \underset{\sim}{\text{r}} A$.

[提示：倒转正文所给的步骤.]

11. 证明断言：对一切矩阵A，$A \underset{r}{\sim} A$。

12. 证明矩阵

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \text{ 和 } \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

不是行-等价的.

13. 证明任意两个不同的 2×4 简化阶梯形矩阵不是行-等价的. 可以把这个结果推广到 $m\times n$ 矩阵。

2.4 高斯化简法：逐次消元

我们现在叙述解线性方程组的高斯化简法，并以下面的方程组为例来说明这一方法：

$$\left.\begin{aligned} -x_2-2x_3+11x_4-3x_5&=15 \\ -2x_1+x_2+2x_3-4x_4+3x_5&=-3 \\ 6x_1-x_2-2x_3-10x_4-3x_5&=-21 \\ -5x_1+2x_2+4x_3-5x_4+6x_5&=-1 \end{aligned}\right\}. \tag{10}$$

这个方程组的增广矩阵是

$$A_b=\begin{pmatrix} 0 & -1 & -2 & 11 & -3 & 15 \\ -2 & 1 & 2 & -4 & 3 & -3 \\ 6 & -1 & -2 & -10 & -3 & -21 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix}. \tag{11}$$

高斯化简法的思想是用一系列行初等变换把 A_b 化为简化阶梯形矩阵. 最终的矩阵确定一个简单的方程组，只要用观察法我们就可以写出这个方程组的通解. 根据我们的理论，这一定也是方程组(10)的通解.

可用以下步骤描述化简过程.

第 1 步　标准化一非零行. 用一个常数乘一非零行使它的首元素变为 1. 例如，用乘 $-\frac{1}{2}$ 来标准化 A_b 的第二行，A_b 的第二行变为

$$\left(1 \quad -\frac{1}{2} \quad -1 \quad 2 \quad -\frac{3}{2} \quad \frac{3}{2}\right).$$

第 1 步需要一个类型II的变换.

第 2 步　清理一列. 用一系列类型III的变换，把一列元素除了一个元素之外全化为 0. 具体地说，要清理(a_{ij})的第 s 列，在第 s 列中选一个非零元素，比如是 a_{rs}，进一步对一切 $i\neq r$，把第 r 行的 $-a_{is}/a_{rs}$ 倍加到第 i 行上. 这 rs-元素称为清理过程的主元. 这一步在相应方程组上的作用是，除了第 r 个方程之外从一切方程消去 x_s.

我们用下表所示的步骤把 A_b 变为简化阶梯形.

表 3

化简过程	矩阵
	$A_b=\begin{pmatrix} 0 & -1 & -2 & 11 & -3 & 15 \\ -2 & 1 & 2 & -4 & 3 & -3 \\ 6 & -1 & -2 & -10 & -3 & -21 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix}$
标准化第一行： 用 -1 乘第一行.	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ -2 & 1 & 2 & -4 & 3 & -3 \\ 6 & -1 & -2 & -10 & -3 & -21 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix}$
用第一行首元素作为主元清理第二列： （i）把第一行的 -1 倍加到第二行上；	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ -2 & 0 & 0 & 7 & 0 & 12 \\ 6 & -1 & -2 & -10 & -3 & -21 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix}$
（ii）把第一行加到第三行上；	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ -2 & 0 & 0 & 7 & 0 & 12 \\ 6 & 0 & 0 & -21 & 0 & -36 \\ -5 & 2 & 4 & -5 & 6 & -1 \end{pmatrix}$
（iii）把第一行的 -2 倍加到第四行上.	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ -2 & 0 & 0 & 7 & 0 & 12 \\ 6 & 0 & 0 & -21 & 0 & -36 \\ -5 & 0 & 0 & 17 & 0 & 29 \end{pmatrix}$

续表

化简过程	矩阵
标准化第二行： 用 $-\frac{1}{2}$ 乘第二行。	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ 1 & 0 & 0 & -\frac{7}{2} & 0 & -6 \\ 6 & 0 & 0 & -21 & 0 & -36 \\ -5 & 0 & 0 & 17 & 0 & 29 \end{pmatrix}$
用第二行首元素作主元，清理第一列： （i）把第二行的 -6 倍加到第三行上；	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ 1 & 0 & 0 & -\frac{7}{2} & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -5 & 0 & 0 & 17 & 0 & 29 \end{pmatrix}$
（ii）把第二行的 5 倍加到第四行上。	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ 1 & 0 & 0 & -\frac{7}{2} & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & -1 \end{pmatrix}$
标准化第四行： 用 -2 乘第四行。	$\begin{pmatrix} 0 & 1 & 2 & -11 & 3 & -15 \\ 1 & 0 & 0 & -\frac{7}{2} & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{pmatrix}$
用第四行的首元素作主元，清理第四列： （i）把第四行的11倍加到第一行上。	$\begin{pmatrix} 0 & 1 & 2 & 0 & 3 & 7 \\ 1 & 0 & 0 & -\frac{7}{2} & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{pmatrix}$
（ii）把第四行的 $\frac{7}{2}$ 倍加到第二行上。	$\begin{pmatrix} 0 & 1 & 2 & 0 & 3 & 7 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{pmatrix}$

最后我们施行一系列行交换（类型 I 的变换），把最末一个矩

阵化为简化阶梯形．我们看到用

(i) 交换第一和第二行，

(ii) 交换第三和第四行，

我们就得到简化阶梯形矩阵

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 3 & 7 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{12}$$

它行-等价于 A_b 且确定方程组

$$\left.\begin{aligned} x_1 \qquad\qquad\qquad\qquad &= 1 \\ x_2 + 2x_3 \qquad + 3x_5 &= 7 \\ x_4 \qquad\qquad &= 2 \end{aligned}\right\}. \tag{13}$$

因为方程组(13)和(10)是等价的，所以它们有相同的通解．这通解是

$$\left.\begin{aligned} & x_1 = 1 \\ & x_4 = 2 \\ & x_3 = p \ (\text{任意}) \\ & x_5 = q \ (\text{任意}) \\ & x_2 = 7 - 2p - 3q \end{aligned}\right\}. \tag{14}$$

这称为解的二参数系，参数是任意常数 p 和 q.

化简过程的总结

把一个矩阵化简为行-等价的简化阶梯形矩阵：

(i) 标准化第一个非零行，把这一行的首元素作为主元，清理包含这个首元素的列；

(ii) 对第二个非零行进行同样的化简，如此继续下去；

(iii) 最后，交换某些行得到阶梯形.

显然用这方法能把任何矩阵化简为行-等价的简化阶梯形矩

阵. 从这个矩阵我们能看出由第一个矩阵所确定方程组的通解.

相容的条件

例 7

简化阶梯形矩阵

$$\begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

所确定的方程组是

$$\begin{aligned} x_1+3\,x_2 \qquad\qquad &=0 \\ x_3+2\,x_4 &=0 \\ 0 &=1, \end{aligned}$$

它显然是不相容的.

如果简化阶梯形矩阵的最后一列(常数列)包含一行的首元素,那末我们总得到一个不相容方程组. 反之,如果最后一列不包含一行的首元素,则我们能写出解,因此方程组相容. 于是我们有下列定理.

定理 II 一个方程组是相容的充要条件是:与它的增广矩阵行-等价的简化阶梯形矩阵的最后一列没有任何行的首元素.

自由变量

在等价方程组(10)和(13)的解(14)里,我们能自由选择 x_3 和 x_5, 则剩下的未知元就被决定或被约束了. 我们说 x_3 和 x_5 是自由变量. 方程组(13)的简化阶梯形矩阵(12)有三个非零行. 每一个非零行约束一个变量. 自由变量的数目等于变量总数和约束变量数的差. 以下定理是这个例子的一般化叙述.

定理 III 如果一个 n 元相容方程组对应的简化阶梯形矩阵有 r 个非零行,那末方程组的通解包含 r 个约束变量和 $n-r$ 个自由变量.

例 8 简化阶梯形矩阵

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 4 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & -5 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -3 \end{pmatrix}$$

确定相容方程组

$$\begin{aligned} x_1+2x_2 \qquad\qquad\qquad +4x_6 &= 2 \\ x_3 \quad +3x_5 \qquad\qquad &= 1 \\ x_4+2x_5-5x_6 &= 3 \\ x_7 &= -3. \end{aligned}$$

这里有七个未知元、四个方程和三个自由变量．我们可以这样选自由变量，使它们不作为一个方程的首项出现，那就是 x_2，x_5 和 x_6．通解是

$$\begin{aligned} &x_2=p, \\ &x_5=q, \\ &x_6=r \ (p,q,r \text{ 任意}), \\ &x_1=2-2p-4r, \\ &x_3=1-3q, \\ &x_4=3-2q+5r, \\ &x_7=-3. \end{aligned}$$

另一种方案，我们可以选 x_3 作为自由变量．此时，由于 $x_5=\frac{1}{3}(1-x_3)$ 约束了 x_5，恰好仍有三个自由变量．然而我们绝不能选 x_7 为自由变量，因为它必须取值 -3．

一个方便的可循规则是：只要阶梯形矩阵的第 i 列不包含一行的首元素，就选 x_i 作为一个自由变量．

我们用一个以后将需要的定理来结束这一节．我们知道齐次方程组是相容的．它具有平凡解 $x_i=0$，对一切 i．以下定理给出了确保齐次方程组非平凡解存在的一个条件．

定理 IV n 元 m 个线性方程的齐次方程组，如果 $m<n$，则必有非平凡解.

证明 假设方程组是

$$\left.\begin{array}{l} a_{11}x_1+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+\cdots+a_{2n}x_n=0 \\ \cdots\quad\cdots\quad\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=0 \end{array}\right\}. \tag{15}$$

增广矩阵

$$A_b=\begin{pmatrix} a_{11} & \cdots & a_{1n} & 0 \\ a_{21} & \cdots & a_{2n} & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} & 0 \end{pmatrix}$$

确定了这方程组.

现在 A_b 行-等价于一个简化阶梯形矩阵 R，我们能用化简方法求出 R。假设 R 有 r 个非零行，那末 $r\leqslant m<n$. 因此根据定理 III，由 R 确定方程组的通解包含 $n-r>0$ 个自由变量. 所以这方程组，以及等价的方程组(15)，有非零解.

练　习

14．用高斯化简法解前面几小节中给出的所有线性方程组.

15．解以下线性方程组

(i)
$$\begin{aligned} x_1-\ x_2-\ x_3+\ x_4&=2 \\ x_1+2x_2-\ x_3+3x_4&=9 \\ 2x_1-5x_2-2x_3\qquad&=-3; \end{aligned}$$

(ii)
$$\begin{aligned} x_1+x_2\qquad&=v_1 \\ x_1\qquad+x_3&=v_2 \\ x_2+x_3&=v_3, \end{aligned}$$

这里 v_1,v_2,v_3 是常数;

(iii)
$$3x_1+7x_2+14x_3+9x_4+30x_5=7$$

$$
\begin{aligned}
2x_1+5x_2+10x_3+6x_4+21x_5&=5\\
3x_1+5x_2+10x_3+8x_4+23x_5&=6\\
2x_1+2x_2+4x_3+5x_4+11x_5&=3.
\end{aligned}
$$

[注意：如果我们用乘$\frac{1}{3}$标准化第一行，便产生了许多分数．这将导致非常麻烦的计算．以类型 III 的变换"把第二行的-1倍加到第一行上"开始化简过程，也许是最好的方法．]

16．可以把高斯化简法用在复系数线性方程组上．解

$$
\begin{aligned}
(1+i)x_1+(i-1)x_2+(2+2i)x_3&=4+6i\\
ix_1-x_2+(1+2i)x_3&=1+5i.
\end{aligned}
$$

17．证明以下方程组有唯一解：

$$
\begin{aligned}
7x_1+x_2+x_3-3x_4&=11\\
x_1+x_2-x_3+3x_4&=-3\\
5x_1+7x_2-x_3+8x_4&=1\\
-x_1-2x_2+3x_3-x_4&=11.
\end{aligned}
$$

再求出前三个方程所构成方程组的通解．

18．证明定理 IV 的以下推广：一个 n 元 m 个方程的相容线性方程组（不一定齐次），如果 $m<n$，则必有无数解．

[提示：解构成一个$(n-r)$参数系，这里 r 是用化简方法所得简化阶梯形矩阵非零行的数目．]

19．n 元 m 个方程的线性方程组，若$m\geqslant n$，它可能不相容，也可能相容；如果相容，可能有唯一解，也可能有无数解．对 $m=n=3$，再对 $m=3, n=2$ 举例说明这个断言．

20．证明以下方程组不相容：

$$
\begin{aligned}
x_1+x_2+3x_3&=4\\
2x_1+3x_2+8x_3-x_4&=12\\
2x_1+x_2+4x_3+x_4&=5.
\end{aligned}
$$

[注意：一旦得到一行(0 0 0 0 1)就可以停止化简过程．这证明了不相容性．]

21．利用行初等变换，把以下矩阵

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & -1 \\ 0 & 1 & 3 \\ -2 & -1 & 1 \end{pmatrix}$$

化为简化阶梯形。

22. 当 $m>n$，证明 $m\times n$ 矩阵行-等价于一个有零行的简化阶梯形矩阵。(用练习 21 说明这一点。)

2.5 舍入误差

现在我们简略考察解线性方程组的实际计算方面的问题。高速计算机的出现，使短时间内解含大量方程和大量未知元的线性方程组成为可能。然而必须记住，计算机容量有限，因而所处理的一切数是舍入到有限小数位数的。例如，分数$\frac{2}{3}$的精确小数展开是 0.666…有无数个 6。一部计算机把$\frac{2}{3}$记作 0.6…66（把精确值截尾）或记作 0.6…67(把精确值舍入)，一直到计算机所容许的那么多 6。因此，即使最大的计算机也产生不出完全准确的解。确实，有时所得出的解和精确解之间的差异是不容忽视的。现在我们举例说明这一情况。

例 9 一个病态方程组

考察方程组

$$\begin{aligned} x_1 + \qquad x_2 &= -2 \\ x_1 + 1.015x_2 &= 1. \end{aligned} \tag{16}$$

其(唯一) 解是 $x_1=-202, x_2=200$.

假定用一部最多只能处理三个有效数字的计算机来解这个方程组。这时舍入 1.015 为 1.01 或 1.02。方程组

$$\begin{aligned} x_1 + \qquad x_2 &= -2 \\ x_1 + 1.01x_2 &= 1 \end{aligned} \tag{17}$$

有解 $x_1=-302$, $x_2=300$. 方程组

$$\begin{aligned} x_1+\quad\ \ x_2&=-2 \\ x_1+1.02x_2&=1 \end{aligned} \tag{18}$$

有解 $x_1=-152$, $x_2=150$.这些解和方程组(16)的解很少或根本没有相似之处. 因此方程(16) 系数的一个小变化就对它的解产生了很大的影响. 称这种方程组为病态的. 把方程组(16)写为

$$\begin{aligned} x_1+\qquad\quad x_2&=-2 \\ x_1+(1+\varepsilon)x_2&=1, \end{aligned} \tag{19}$$

这里 $\varepsilon=0.015$,我们就可以看出病态的解释. 方程组(19) 的解是

$$x_1=-\frac{3+2\varepsilon}{\varepsilon}, \qquad x_2=\frac{3}{\varepsilon},$$

对 ε 的小变化这解是很敏感的.

例 10 一个良态方程组

把病态方程(16) 和以下只改变(16) 一个符号所得方程组相比较:

$$\begin{aligned} x_1+\qquad\ \ x_2&=-2 \\ x-1.015x_2&=1, \end{aligned} \tag{20}$$

它的四个有效数字的通解是

$$x_1=-0.5111, \quad x_2=-1.489. \tag{21}$$

由方程(20) 舍入-1.015 为-1.01 所得方程组有解

$$x_1=-0.5075, \quad x_2=-1.492.$$

如果舍入-1.015 为-1.02,则新方程组有解

$$x_1=-0.5148, \quad x_2=-1.485.$$

这些解和解(21) 相一致到两个有效数字. 方程组(20) 系数的一个小变化只产生了它解的一个小变化. 这种方程组称为良态的. 如果我们把方程组写为

$$
\left.\begin{array}{l}
x_1 + \qquad x_2 = -2 \\
x_1 - (1+\varepsilon)x_2 = 1
\end{array}\right\} \tag{22}
$$

这里 $\varepsilon = 0.015$，那末方程组(20)的稳定性是明显的．方程组(22)的解是

$$
x_1 = -\frac{1+2\varepsilon}{2+\varepsilon}, \quad x_2 = -\frac{3}{2+\varepsilon}.
$$

这里不象例 9，分母不是小的．因此 ε 值的小变化只引起方程组(22)解的小变化．

主元的选择

解线性方程的高斯化简法只适合于理论目的．在实用中改进它的理由之一是为了能够尽可能地避免病态方程．这就需要小心选择所有第 III 种变换的主元．我们用一个简单的例子来说明．

例 11

方程组

$$
\left.\begin{array}{l}
x_1 + \qquad x_2 + \qquad x_3 = 2 \\
x_1 + 0.0001x_2 + 0.0001x_3 = 0.0001 \\
x_1 + 0.0001x_2 - 0.0001x_3 = -0.0001
\end{array}\right\} \tag{23}
$$

的解是

$$
x_1 = -\frac{1}{9999}, \quad x_2 = 1\frac{1}{9999}, \quad x_3 = 1. \tag{24}
$$

如果我们把所有系数舍入为三个有效数字，一开始从第二和第三个方程消去 x_1(即以 11-元素作为主元)，再把所得解和解(24)相比较．方程组(23)后两个方程变为

$$
\left.\begin{array}{l}
0.9999\,x_2 + 0.9999\,x_3 = 1.9999 \\
0.9999\,x_2 + 1.0001\,x_3 = 2.0001
\end{array}\right\} \tag{25}
$$

它们都舍入为

$$
x_2 + \qquad x_3 = 2. \tag{26}
$$

于是得到错误结论，方程组(23)有无数解

$$x_3=h\ (\text{任意}),\quad x_2=2-k,\quad x_1=0.$$

通过舍入,错误发生了,可见方程组(25)是病态的.

改变主元避免了上述困难. 如果我们从方程组(23)的第一和第三个方程消去 x_1(用 12-元素作为主元), 得到

$$\begin{aligned} 0.9999\,x_2+0.9999\,x_3&=1.9999 \\ 0.0002\,x_3&=0.0002. \end{aligned} \tag{27}$$

(27)中第一个方程舍入后为

$$x_2+x_3=2.$$

第二个方程保持不变, 因为系数 0.0002 ($=2\times10^{-4}$)只包含一个有效数字. 因此我们得到近似解

$$x_3=1,\quad x_2=1,\quad x_1=0,$$

它是很靠近准确解(24)的.

在上例的解题过程中,固定了有效数字的个数,但没有仅因某系数看起来很小而丢掉它. 事实上,可以用 10 的幂乘(27)的第二个方程,非常简单地改变这方程系数的大小,又不影响有效数字的个数. 于是我们看到这个方程等价于 $2x_3=2$.

许多计算机都设计成允许这种小数点移动. 我们称这些计算机是采用浮点算法的. 当然, 在某些情形下定点算法更为合理.

关于线性方程组数值解的进一步的资料, 可以在下列参考书中找到.

1. B. Noble, *Numerical Methods 1. Iteration, Programming and Algebraic Equations*, University Mathematical Text, Oliver & Boyd, 1964.

2. W. G. Bickley & R. S. H. G. Thompson, *Matrices, Their Meaning and Manipulation*, English University Press, 1964.

3. B. Noble, *Notes on the Numerical Solution of*

Simultaneous Linear Algebraic Equations, The University of Wisconsin, M. R. C. Report No. 644AD, U. S. Army, 1966.

练　　习

23. 考察系数 1.985 一个小变化来证明方程组

$$x+\quad 2y=0$$
$$x+1.985y=-1$$

是病态的。

24. 证明方程组

$$x-\quad 2y=0$$
$$x+1.985y=-1$$

是良态的。

25. 证明，当 $a_{11}a_{22}-a_{12}a_{21}$ 非零且和系数 a_{11} 和 a_{12} 相比为小的时候，方程组

$$a_{11}x_1+a_{12}x_2=0$$
$$a_{21}x_1+a_{22}x_2=-1$$

是病态的。对照这条规则检验练习 23 和 24。

26. 证明方程组

$$0.0001\,x_1+0.0001\,x_2=0$$
$$x_1+\quad 2x_2=-1$$

是良态的。这里 $a_{11}a_{22}-a_{12}a_{21}$ 似乎是"小的"。然而，它和系数 a_{11} 和 a_{12} 是同一数量级的，因此，这个方程是良态的。除了直接比较二数之外，术语"小"和"大"就没有多少意义了。

27. 如果 $a_{11}a_{22}-a_{12}a_{21}=0$，证明练习 25 的方程组或者不相容，或者有无数解。(表达式 $a_{11}a_{22}-a_{12}a_{21}$ 称为系数矩阵的行列式。后面有一章中我们将研究行列式和它们在线性方程组中的应用。)

28. 在例 11 方程组(23)中，置 $y_1=10^4x_1$。解所得到的 y_1, x_2, x_3 的方程组。首先在第二，第三个方程中约去 10^{-4}。然后可以看出，适合方程组的 y_1, x_2, x_3 的值都是同一数量级的。(称形如 $y_i=10^k x_i$ 的代换改变了方程组的比例。)

第三章　向量空间的理论

在这一章里，我们继续向量空间的研究．我们要引进线性组合，线性无关和基的概念，并用这些概念来讨论向量空间的维数．

3.1　线性组合：向量空间的生成元

设$\mathscr{V}$是向量空间．给了$\mathscr{V}$中的向量$\mathbf{v}_1,\cdots,\mathbf{v}_r$，以及纯量$c_1,\cdots,c_r$，向量

$$c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r$$

称为$\mathbf{v}_1,\cdots,\mathbf{v}_r$的线性组合．

例 1

在$\mathscr{R}^3$中，设

$$\mathbf{v}=\left(1,-1,\frac{1}{8}\right),\quad \mathbf{w}=(-2,0,-1),\quad \mathbf{s}=\left(1,-2,-\frac{1}{4}\right).$$

则$\mathbf{s}$是$\mathbf{v}$与$\mathbf{w}$的线性组合，因为

$$\mathbf{s}=2\mathbf{v}+\frac{1}{2}\mathbf{w}.$$

此外，$\mathbf{v}$也是$\mathbf{s}$和$\mathbf{w}$的线性组合，因为

$$\mathbf{v}=\frac{1}{2}\mathbf{s}+\left(-\frac{1}{4}\right)\mathbf{w}.$$

例 2

定义函数

$$y_1(t)=\sin\omega t,\quad y_2(t)=\cos\omega t,$$

函数y_1和y_2的所有线性组合构成微分方程

$$\frac{d^2y}{dt^2}+\omega^2y=0\quad(\omega\neq0)$$

(1.3节例3)的通解．

向量空间的生成元

向量空间$\mathscr{V}$的向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 的一切可能线性组合的集合构成$\mathscr{V}$的子空间$\mathscr{S}$. 因为对$\mathscr{S}$中的任意向量

$$\mathbf{s}=c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r$$

与

$$\mathbf{t}=d_1\mathbf{v}_1+\cdots+d_r\mathbf{v}_r,$$

$\mathscr{S}$也包含向量

$$\mathbf{s}+\mathbf{t}=(c_1+d_1)\mathbf{v}_1+\cdots+(c_r+d_r)\mathbf{v}_r$$

与

$$k\mathbf{s}=(kc_1)\mathbf{v}_1+\cdots+(kc_r)\mathbf{v}_r,$$

对一切纯量 k. 由 1.3 节练习 10 即得出结论.

如果$\mathscr{S}$的每个向量可表作 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 的线性组合，我们称$\mathscr{V}$的子空间$\mathscr{S}$由 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 生成，称 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 为$\mathscr{S}$的生成元. 特别地，如果$\mathscr{S}=\mathscr{V}$则$\mathscr{V}$是由$\mathbf{v}_1,\cdots,\mathbf{v}_r$生成的. 一个可由有限个向量生成的向量空间叫作有限生成的向量空间.

例 3

微分方程

$$\frac{d^2y}{dt^2}+\omega^2y=0$$

所有解的向量空间由上述例 2 中定义的函数 y_1 和 y_2 生成.

例 4 生成向量空间$\mathscr{S}$的向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 本身属于$\mathscr{S}$.

每个 $\mathbf{v}_i$ $(i=1,\cdots,r)$可表示为 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 的线性组合. 例如

$$\mathbf{v}_2=0\,\mathbf{v}_1+1\,\mathbf{v}_2+0\,\mathbf{v}_3+\cdots+0\,\mathbf{v}_r.$$

例 5 向量空间 $\mathscr{R}^3$ 由下列三个向量生成

$$\mathbf{e}_1=(1,0,0),\quad \mathbf{e}_2=(0,1,0),\quad \mathbf{e}_3=(0,0,1).$$

这是因为 $\mathscr{R}^3$ 中任意一个向量 $\mathbf{v}=(v_1,v_2,v_3)$可表作 $\mathbf{e}_1,\mathbf{e}_2$ 和 $\mathbf{e}_3$ 的线性组合：

$$\mathbf{v}=v_1\mathbf{e}_1+v_2\mathbf{e}_2+v_3\mathbf{e}_3.$$

例 6 还有许多其它生成 $\mathscr{R}^3$ 的方法。例如，向量 $\mathbf{b}_1=(1,1,0)$，$\mathbf{b}_2=(1,0,1)$，$\mathbf{b}_3=(0,1,1)$ 生成 $\mathscr{R}^3$.

我们必须证明 $\mathbf{v}=(v_1,v_2,v_3)$ 可表作 $\mathbf{b}_1,\mathbf{b}_2$ 和 $\mathbf{b}_3$ 的线性组合. 假设存在纯量 x_1,x_2 和 x_3 使

$$\mathbf{v}=x_1\mathbf{b}_1+x_2\mathbf{b}_2+x_3\mathbf{b}_3. \tag{1}$$

则

$$\begin{aligned}(v_1,v_2,v_3)&=(x_1,x_1,0)+(x_2,0,x_2)+(0,x_3,x_3)\\&=(x_1+x_2,x_1+x_3,x_2+x_3).\end{aligned}$$

因此我们要求 x_1,x_2 和 x_3 适合线性方程组

$$\begin{aligned}x_1+x_2\qquad&=v_1\\x_1\qquad+x_3&=v_2\\x_2+x_3&=v_3.\end{aligned}$$

这方程组的解为

$$x_1=\frac{1}{2}(v_1+v_2-v_3)$$

$$x_2=\frac{1}{2}(v_1-v_2+v_3)$$

$$x_3=\frac{1}{2}(-v_1+v_2+v_3).$$

在 (1) 中如果这样选择 x_1,x_2 和 x_3，向量 $\mathbf{v}=(v_1,v_2,v_3)$ 就表成了 $\mathbf{b}_1,\mathbf{b}_2$ 和 $\mathbf{b}_3$ 的一个线性组合.

例 7 不应该由前面两例推断 $\mathscr{R}^3$ 的任意三个向量生成整个三-空间. 例如，向量

$$\mathbf{v}=(1,-1,\frac{1}{8}),\quad \mathbf{w}=(-2,0,-1),\quad \mathbf{s}=(1,-2,-\frac{1}{4})$$

生成 $\mathscr{R}^3$ 的一个真子空间 $\mathscr{S}$.

因为

$$\mathbf{s}=2\mathbf{v}+\frac{1}{2}\mathbf{w},$$

所以 $\mathscr{S}$ 中每个向量可表作 $\mathbf{v}$ 和 $\mathbf{w}$ 的线性组合：

$$x_1\mathbf{v}+x_2\mathbf{w}+x_3\mathbf{s}=(x_1+2x_3)\mathbf{v}+(x_2+\frac{1}{2}x_3)\mathbf{w}.$$

为了证明 $\mathscr{S}$ 是 $\mathscr{R}^3$ 的真子空间，我们只需在 $\mathscr{R}^3$ 中拿出一个不能表作 $\mathbf{v}$ 和 $\mathbf{w}$ 线性组合的向量即可. (1,0,0) 就是这样的向量. 假设

$$(1,0,0)=y_1(1,-1,\frac{1}{8})+y_2(-2,0,-1).$$

则 y_1, y_2 必须适合方程组

$$\begin{aligned} y_1-2y_2&=1\\ -y_1\qquad&=0\\ \frac{1}{8}y_1-\ y_2&=0. \end{aligned}$$

这方程组不相容. 因此 $\mathscr{S}$ 是 $\mathscr{R}^3$ 的真子空间.

几何上可以这样解释这个例子：当我们把 $\mathscr{R}^3$ 的向量看作关于某个坐标系原点的位置向量时，$\mathscr{S}$ 中每个向量都在包含 $\mathbf{v}$ 和 $\mathbf{w}$ 的平面中，而向量(1,0,0)在这平面之外.

练　　习

1. $\mathscr{R}^2$ 中，设

$$\mathbf{v}=(1,1),\quad \mathbf{w}=(2,-1).$$

把下列向量表作 $\mathbf{v}$ 和 $\mathbf{w}$ 的一个线性组合：

(i) (8,−1)；

(ii) (0,−3)；

(iii) (−6,3).

这些向量中有能用不只一种方法表作 $\mathbf{v}$ 和 $\mathbf{w}$ 的线性组合的向量吗？用图说明.

2. 以下哪些向量集合生成 $\mathscr{R}^3$？

(i) (1,1,0),(2,2,0),(0,0,1)；

(ii) (1,−1,0),(1,0,−1),(0,1,−1)；

(iii) $(1,1,1),(1,-1,1),(1,-1,-1)$;

(iv) $(1,1,0),(1,0,0),(0,1,0),(0,0,1)$.

3. $\mathscr{R}^4$ 中，设

$$\mathbf{a}=(1,1,0,1),\quad \mathbf{b}=(1,0,1,1),\quad \mathbf{c}=(0,1,-1,0),\quad \mathbf{d}=(1,1,1,0).$$

证明 $\mathbf{a}$ 是 $\mathbf{b},\mathbf{c}$ 与 $\mathbf{d}$ 的线性组合.

[提示：找出纯量 x_1,x_2 和 x_3，使 $\mathbf{a}=x_1\mathbf{b}+x_2\mathbf{c}+x_3\mathbf{d}$.]

再证明 $\mathbf{d}$ 不是 $\mathbf{a},\mathbf{b}$ 和 $\mathbf{c}$ 的线性组合.

4. 证明 $\mathscr{R}^4$ 不能由练习 3 的向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 和 $\mathbf{d}$ 生成.

5. 求一个向量，使它和 $\mathbf{a},\mathbf{b},\mathbf{c}$ 一起生成 $\mathscr{R}^4$.

6. 如果 $a_{11}a_{22}-a_{12}a_{21}\neq 0$，证明

$$\mathbf{b}_1=(a_{11},a_{21})\quad 和\quad \mathbf{b}_2=(a_{12},a_{22})$$

生成 $\mathscr{R}^2$.

[提示：应用例 6 的方法.]

再证明其逆：如果 $a_{11}a_{22}-a_{12}a_{21}=0$，则 $\mathbf{b}_1$ 和 $\mathbf{b}_2$ 不生成 $\mathscr{R}^2$.

[提示：证明要末 $\mathbf{b}_1=\mathbf{0}$，要末 $\mathbf{b}_1\neq\mathbf{0}$ 且对某个纯量 k，$\mathbf{b}_2=k\mathbf{b}_1$. 无论哪种情形，都可找到一个向量，它不是 $\mathbf{b}_1$ 和 $\mathbf{b}_2$ 的线性组合.]

用数值例子说明这些结果.

7. 以下哪些向量集合生成 $\mathscr{C}^3$?

(i) $(1,i,i),(i,1,i),(i,i,1)$.

(ii) $(1,i,i),(i,1,1),(0,1,1)$.

3.2 线性相关和线性无关

我们称向量空间向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 是线性相关的，如果它们满足一个非平凡线性关系，即有一组不全为零的纯量 $c_1,\cdots,c_r$，使

$$c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r=\mathbf{0}. \tag{2}$$

如果 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 所满足的线性关系只是平凡的线性关系

$$0\,\mathbf{v}_1+\cdots+0\,\mathbf{v}_r=\mathbf{0},$$

则我们称 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 是线性无关的.

例 8　如果向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 中有一个是零向量，则它们线性

相关.

假设 $\mathbf{v}_r=\mathbf{0}$. 则 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 适合非平凡关系

$$0\,\mathbf{v}_1+\cdots+0\,\mathbf{v}_{r-1}+c_r\mathbf{v}_r=\mathbf{0},$$

这里 $c_r\neq 0$.

例 9 向量

$$\mathbf{v}_1=(0,2,a,b,c,d),\quad \mathbf{v}_2=(0,0,0,4,e,f),$$
$$\mathbf{v}_3=(0,0,0,0,-1,g)$$

是线性无关的.

假设

$$\mathbf{w}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}.$$

因为 $\mathbf{w}$ 第二个元素是 0,所以我们必有 $c_1=0$. 因为 $\mathbf{w}$ 第四个元素是 0,所以我们必有 $c_2=0$. 最后我们有 $c_3=0$. 这样 $\mathbf{v}_1,\mathbf{v}_2$ 和 $\mathbf{v}_3$ 所适合的线性关系只有平凡关系了.

我们可把例 9 推广如下:设 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 是 $\mathscr{R}^n$ 中非零向量,它们的首元素出现得一个比一个向右,那末 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 线性无关.

例 10 让我们回忆 3.1 节的例题.

(例 1) 向量 $\mathbf{v},\mathbf{w}$ 和 $\mathbf{s}$ 是线性相关的. 它们适合非平凡关系

$$2\,\mathbf{v}+\frac{1}{2}\mathbf{w}+(-1)\mathbf{s}=\mathbf{0}.$$

(例 2) 函数 y_1 和 y_2 线性无关.假定

$$c_1\sin\omega t+c_2\cos\omega t=0 \tag{3}$$

对一切 t.(3)中置 $t=0$ 我们得到 $c_2=0$. 因此也有 $c_1=0$. 于是 y_1 和 y_2 所适合的线性关系只有平凡关系.

(例 5) 向量

$$\mathbf{e}_1=(1,0,0),\quad \mathbf{e}_2=(0,1,0),\quad \mathbf{e}_3=(0,0,1)$$

是线性无关的. 因为首元素出现得一个比一个向右.

(例 6) 向量$(1,1,0)$, $(1,0,1)$和$(0,1,1)$线性无关. 我们将在后面(例 12)用以下线性相(无)关判别法来证明这一点.

$\mathscr{R}^n$ 中线性相关判别法

设

$$\mathbf{v}_1=(v_{11},v_{12},\cdots,v_{1n}),\quad \mathbf{v}_2=(v_{21},v_{22},\cdots,v_{2n}),$$
$$\cdots,\quad \mathbf{v}_r=(v_{r1},v_{r2},\cdots,v_{rn})$$

是 $\mathscr{R}^n$ 的 r 个向量．设 A 是行为 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 的 $r\times n$ 矩阵：

$$A=\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ v_{r1} & v_{r2} & & v_{rn} \end{pmatrix}.$$

用行初等变换把 A 化简为阶梯形 B．若 B 有零行，则 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 线性相关．否则 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 线性无关．

为了证明这个判别法，我们需要一个引理．

引理 I (a) 已知 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 线性相关，则

$$\mathbf{v}_1,\cdots,c\mathbf{v}_i,\cdots,\mathbf{v}_r \quad (c\neq 0) \tag{4}$$

和

$$\mathbf{v}_1,\cdots,\mathbf{v}_i+k\mathbf{v}_j,\cdots,\mathbf{v}_r \quad (i\neq j) \tag{5}$$

也是线性相关的．

(b) 已知 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 线性无关，则向量(4)和(5)也是线性无关的．

证明 (a) 假设 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 线性相关．则它们适合一个非平凡关系

$$c_1\mathbf{v}_1+\cdots+c_i\mathbf{v}_i+\cdots+c_r\mathbf{v}_r=\mathbf{0}. \tag{6}$$

这就得出

$$c_1\mathbf{v}_1+\cdots+\frac{c_i}{c}(c\mathbf{v}_i)+\cdots+c_r\mathbf{v}_r=\mathbf{0}$$

是 $\mathbf{v}_1,\cdots,c\mathbf{v}_i,\cdots,\mathbf{v}_r$ 的一个非平凡关系．因此 $\mathbf{v}_1,\cdots,c\mathbf{v}_i,\cdots,\mathbf{v}_r$ 是线性相关的．

从关系式(6)我们还能求得

$$c_1\mathbf{v}_1+\cdots+c_i(\mathbf{v}_i+k\mathbf{v}_j)+\cdots+(c_j-kc_i)\mathbf{v}_j+\cdots+c_r\mathbf{v}_r=\mathbf{0}$$

这是 $\mathbf{v}_1,\cdots,\mathbf{v}_i+k\mathbf{v}_j,\cdots,\mathbf{v}_r$ 的一个非平凡关系．因为若每个系数 $c_1,\cdots,c_i,\cdots,c_j-kc_i,\cdots,c_r$ 等于零，则(6)中每个系数也等于零．这和我们的假设矛盾．因此 $\mathbf{v}_1,\cdots,\mathbf{v}_i+k\mathbf{v}_j,\cdots,\mathbf{v}_r$ 线性相关．

证明 (b) 存在 $\mathbf{v}_1,\cdots,c\,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 或 $\mathbf{v}_1,\cdots,\mathbf{v}_i+k\mathbf{v}_j,\cdots,\mathbf{v}_r$ 的一个非平凡关系导致一个 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 的非平凡关系．因此如果 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_r$ 线性无关，则向量(4)和(5)也都线性无关．

推论 I 设 A 和 B 是 $r\times n$ 行-等价矩阵．若 A 和 B 的行向量依次是 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 和 $\mathbf{w}_1,\cdots,\mathbf{w}_r$，则 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 和 $\mathbf{w}_1,\cdots,\mathbf{w}_r$ 或者都线性相关，或者都线性无关．

证明 引理 I 证明了第 II 和第 III 种行初等变换保持行向量的线性相关或线性无关．对第 I 种变换这也显然是成立的．

现在我们可以证明 $\mathscr{R}^n$ 中线性相关的判别法了．假设 B 有零行，则 B 的行是线性相关的（例 8）．由推论 I，A 的行也线性相关．

反之，假设 B 没有零行．因为 B 是阶梯形的，所以它行的首元素出现得一个比一个向右，因此 B 的行线性无关(例 9)．由推论 I，A 的行也线性无关．

例 11 判别向量

$$\mathbf{v}_1=\left(0,2,5,1,\frac{1}{2}\right)$$

$$\mathbf{v}_2=\left(2,2,-\frac{1}{2},1,1\right)$$

$$\mathbf{v}_3=(-8,-1,12,4,-3)$$

$$\mathbf{v}_4=(4,3,-1,0,2)$$

的线性相关性．

设 A 是行为 $\mathbf{v}_1$，$\cdots$，$\mathbf{v}_4$ 的矩阵．我们用行初等变换把 A 化简为阶梯形．如果我们把行向量加上标记，那末所施行的行初等变换就能清楚地看出：

$$A=\begin{pmatrix} 0 & 2 & 5 & 1 & \frac{1}{2} \\ 2 & 2 & -\frac{1}{2} & 1 & 1 \\ -8 & -1 & 12 & 4 & -3 \\ 4 & 3 & -1 & 0 & 2 \end{pmatrix}\begin{matrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \end{matrix}$$

$$\overset{r}{\sim}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 2 & 5 & 1 & \frac{1}{2} \\ -8 & -1 & 12 & 4 & -3 \\ 4 & 3 & -1 & 0 & 2 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \end{matrix}$$

$$\overset{r}{\sim}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 2 & 5 & 1 & \frac{1}{2} \\ 0 & 7 & 10 & 8 & 1 \\ 4 & 3 & -1 & 0 & 2 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1 \\ \mathbf{v}_3+4\,\mathbf{v}_2 \\ \mathbf{v}_4 \end{matrix}$$

$$\overset{r}{\sim}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 2 & 5 & 1 & \frac{1}{2} \\ 0 & 7 & 10 & 8 & 1 \\ 0 & -1 & 0 & -2 & 0 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1 \\ \mathbf{v}_3+4\,\mathbf{v}_2 \\ \mathbf{v}_4-2\,\mathbf{v}_2 \end{matrix}$$

$$\underset{\sim}{r}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 0 & 5 & -3 & \frac{1}{2} \\ 0 & 7 & 10 & 8 & 1 \\ 0 & -1 & 0 & -2 & 0 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1+2(\mathbf{v}_4-2\mathbf{v}_2) \\ \mathbf{v}_3+4\mathbf{v}_2 \\ \mathbf{v}_4-2\mathbf{v}_2 \end{matrix} \qquad (7)$$

$$\underset{\sim}{r}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 0 & 5 & -3 & \frac{1}{2} \\ 0 & 0 & 10 & -6 & 1 \\ 0 & -1 & 0 & -2 & 0 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1-4\mathbf{v}_2+2\mathbf{v}_4 \\ (\mathbf{v}_3+4\mathbf{v}_2)+7(\mathbf{v}_4-2\mathbf{v}_2) \\ \mathbf{v}_4-2\mathbf{v}_2 \end{matrix}$$

$$\underset{\sim}{r}\begin{pmatrix} 2 & 2 & -\frac{1}{2} & 1 & 1 \\ 0 & 0 & 5 & -3 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -2 & 0 \end{pmatrix}\begin{matrix} \mathbf{v}_2 \\ \mathbf{v}_1-4\mathbf{v}_2+2\mathbf{v}_4 \\ \mathbf{w} \\ \mathbf{v}_4-2\mathbf{v}_2 \end{matrix}$$

这里
$$\begin{aligned}\mathbf{w}&=\{\mathbf{v}_3+4\mathbf{v}_2+7(\mathbf{v}_4-2\mathbf{v}_2)\}-2\{\mathbf{v}_1-4\mathbf{v}_2+2\mathbf{v}_4\}\\&=-2\mathbf{v}_1-2\mathbf{v}_2+\mathbf{v}_3+3\mathbf{v}_4.\end{aligned}$$

我们不必再作下去了．因为已经得到一个零行，所以我们断言向量 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 和 $\mathbf{v}_4$ 是线性相关的．因为我们有行向量的标记，所以我们能写出一个非平凡线性关系．它是 $\mathbf{w}=\mathbf{0}$，即

$$-2\mathbf{v}_1-2\mathbf{v}_2+\mathbf{v}_3+3\mathbf{v}_4=\mathbf{0}. \qquad (8)$$

上例中我们选用 42-元素作主元来清理第二列（(7) 式）．我们认为这个主元比 22-元素好，仅仅是因为这样算起来方便．

例 12 判断向量

$$\mathbf{v}_1=(1,1,0),\quad \mathbf{v}_2=(1,0,1),\quad \mathbf{v}_3=(0,1,1)$$

是线性相关的还是线性无关的．

$$A=\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}\begin{matrix}\mathbf{v}_1\\ \mathbf{v}_2\\ \mathbf{v}_3\end{matrix}\qquad \underset{\sim}{r}\begin{pmatrix}1 & 1 & 0\\ 0 & -1 & 1\\ 0 & 1 & 1\end{pmatrix}\begin{matrix}\mathbf{v}_1\\ \mathbf{v}_2-\mathbf{v}_1\\ \mathbf{v}_3\end{matrix}$$

$$\underset{\sim}{r}\begin{pmatrix}1 & 1 & 0\\ 0 & -1 & 1\\ 0 & & 2\end{pmatrix}\begin{matrix}\mathbf{v}_1\\ \mathbf{v}_2-\mathbf{v}_1\\ \mathbf{v}_3+\mathbf{v}_2-\mathbf{v}_1\end{matrix}$$

这最后的矩阵是阶梯形的(但不是简化阶梯形). 它的每一行都有非零元素. 因此由判别法,向量 $\mathbf{v}_1,\mathbf{v}_2$ 和 $\mathbf{v}_3$ 线性无关.

例 13 假设向量 $\mathbf{v}_1,\cdots\ \mathbf{v}_r$ 线性相关. 证明这些向量中至少有一个可表作其余向量的线性组合.

存在一个非平凡关系

$$c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r=\mathbf{0},$$

其中 $c_i\neq 0$,对某个 i。

因此

$$\mathbf{v}_i=-\left(\frac{c_1}{c_i}\right)\mathbf{v}_1-\cdots-\left(\frac{c_{i-1}}{c_i}\right)\mathbf{v}_{i-1}-\left(\frac{c_{i+1}}{c_i}\right)\mathbf{v}_{i+1}-\cdots$$

$$-\left(\frac{c_r}{c_i}\right)\mathbf{v}_r,$$

我们已把 $\mathbf{v}_i$ 表作了 $\mathbf{v}_1,\cdots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\cdots,\mathbf{v}_r$ 的线性组合。

练　习

8. 判别下列 $\mathscr{R}^3$ 向量集合的线性相关性. 如果线性相关，试求一个非平凡关系，再把这个集合的某一个向量表示为其余向量的线性组合.

(i) (1, 1, 0), (2, 2, 0), (0, 0, 1);

(ii) (1, −1, 0), (1, 0, −1), (0, 1, −1);

(iii) (1, 1, 1), (1, −1, 1), (1, −1, −1);

(iv) (1, 1, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1).

9. 证明 $\mathscr{R}^3$ 中任意四个向量线性相关. 一般地证明，如果 $m>n$,则 $\mathscr{R}^n$ 中任意 m 个向量线性相关.

10．判别下列 $\mathscr{C}^3$ 向量集合的线性相关性

(i) $(1,i,i)$，$(i,1,i)$，$(i,i,1)$；

(ii) (i,i,i)，$(i,1,1)$，$(0,1,1)$.

11．用线性相关判别法，再作 3.1 节练习 3．注意，虽然 $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ 线性相关，但向量 $\mathbf{d}$ 不可表作 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的线性组合.

12．如果 $a_{11}a_{22}-a_{12}a_{21}\neq 0$，证明 $\mathbf{b}_1=(a_{11},a_{21})$ 和 $\mathbf{b}_2=(a_{12},a_{22})$ 线性无关.

[提示：注意 a_{11} 和 a_{21} 中至少有一个一定不是零.]

反之正确吗？（和 3.1 节练习 6 相比较.）

13．证明，如果 $\mathbf{v}$ 和 $\mathbf{w}$ 线性相关，则必有

$$\mathbf{w}=\mathbf{0} \quad 或 \quad \mathbf{v}=k\mathbf{w}$$

对某纯量 k.

14．证明如果一个向量空间由 r 个线性相关向量生成，则可选择这些向量中的 $r-1$ 个生成相同的向量空间.（作为数值实例，可见 3.1 节例 7.）

3.3 基和维数

设 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 是向量空间 $\mathscr{V}$ 的向量，使得

(i) $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 生成 $\mathscr{V}$，

(ii) $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 线性无关.

则我们称 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 构成 $\mathscr{V}$ 的基.

例 14 向量

$$\mathbf{e}_1=(1,0,0), \quad \mathbf{e}_2=(0,1,0), \quad \mathbf{e}_3=(0,0,1)$$

构成 $\mathscr{R}^3$ 的基.

向量

$$\mathbf{b}_1=(1,1,0), \quad \mathbf{b}_2=(1,0,1), \quad \mathbf{b}_3=(0,1,1)$$

构成 $\mathscr{R}^3$ 的另一组基.

向量 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 和 $\mathbf{b}_1$ 不构成 $\mathscr{V}$ 的基，因为虽然它们生成 $\mathscr{V}$，但它们不是线性无关的.

向量

$$\mathbf{v}=\left(1,-1,\frac{1}{8}\right),\quad \mathbf{w}=(-2,0,-1),\quad \mathbf{s}=\left(1,-2,-\frac{1}{4}\right)$$

不构成 $\mathscr{R}^3$ 的基.因为它们既不是线性无关的，又不生成 $\mathscr{R}^3$(它们生成一个真子空间).

例 15 证明有限生成向量空间 $\mathscr{V}$ 有有限基.

假设 $\mathscr{V}$ 由 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 生成. 如果这些向量线性无关，则它们构成有限基. 如若不然，我们可以应用 3.2 节练习 14， 得到 $s(<r)$ 个向量,它们生成 $\mathscr{V}$ 并且线性无关.

例 16 设 $\mathscr{V}$ 是 n 个向量生成的向量空间. 证明对 $m>n$，$\mathscr{V}$ 的任意 m 个向量线性相关.

假设 $\mathscr{V}$ 由 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 生成. $\mathscr{V}$ 的任意向量可表作线性组合 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$，且可以方便地记之为 n 元组 $(c_1,\cdots,c_n)$. 可用与 $\mathscr{R}^n$ 中向量相同的方法，对这些 n 元组进行加法和纯量乘法.

设 $\mathbf{b}_1,\cdots,\mathbf{b}_m$ 是 $\mathscr{V}$ 中 m 个向量，则对 $i=1,\cdots,m$，$\mathbf{b}_i$ 可表作 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 的线性组合：

$$\mathbf{b}_i=a_{i1}\mathbf{v}_1+\cdots+a_{in}\mathbf{v}_n=(a_{i1},\cdots,a_{in}).$$

现在,因为 $m>n$，矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

行-等价于一个有零行的阶梯形矩阵(2.4 节练习 22). 因此,根据我们的判别法,向量 $\mathbf{b}_1,\cdots,\mathbf{b}_m$ 是线性相关的.

定理 I 设 $\mathscr{V}$ 是有限生成向量空间，则 $\mathscr{V}$ 的每一组基都有相同个数的向量.

证明 由例 15，$\mathscr{V}$ 有一组有限基. 设 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 是一组基，

其中所含向量个数最少；而 $\mathbf{b}_1,\cdots,\mathbf{b}_m$ 是 $\mathscr{V}$ 的另外任何一组基．由我们 n 的选法，知 $m\geqslant n$．而如果 $m>n$，则由例 16，$\mathbf{b}_1,\cdots,\mathbf{b}_m$ 线性相关并因此不构成 $\mathscr{V}$ 的基．于是有 $m=n$．

维数

设 $\mathscr{V}$ 是有限生成向量空间．$\mathscr{V}$ 的任何一组基中的向量个数称为 $\mathscr{V}$ 的维数．如果这个数是 n，我们称 $\mathscr{V}$ 是一个 n 维向量空间．如果 $\mathscr{V}$ 对某有限数 n 有一组 n 个向量的基，则我们称 $\mathscr{V}$ 是有限维的．

例 17

(a) $\mathscr{R}^n$ 是一个 n 维向量空间．

(b) $\dfrac{d^2y}{dt^2}+\omega^2y=0 \quad (\omega\neq 0)$

的解构成 2 维向量空间．这空间的一组基是 y_1, y_2，这里

$$y_1(t)=\sin\omega t \quad 和 \quad y_2(t)=\cos\omega t.$$

例 18 多项式

设 $a_0,a_1,\cdots,a_n$ 是实数．我们称

$$p(x)=a_0+a_1x+\cdots+a_nx^n$$

为 x 的一个(实)多项式．如果 $a_n\neq 0$，我们称 $p(x)$ 是 n 次的．把一个实数 a_0 当作 0 次多项式．这里我们包括了 $a_0=0$ 这种可能，在这种情形我们有零多项式 $\theta(x)=0$．两个多项式

$$a_0+a_1x_1+\cdots+a_nx^n \quad 和 \quad b_0+b_1x+\cdots+b_mx^m$$

相等当且仅当它们有相同次数且 $a_i=b_i$，对每个 i．

可用自然的方法把两个多项式相加或相乘而得到一个新多项式．

一切次数 $<n$ 的多项式集合构成一个 n 维向量空间 $\mathscr{P}_n$．n 个多项式

$$p_0(x)=1,\ p_1(x)=x,\ \cdots,p_r(x)=x^r,\cdots,\ p_{n-1}(x)=x^{n-1}$$

构成 $\mathscr{P}_n$ 的一组基．

证明　这些多项式生成 $\mathscr{P}_n$，因为任意一个次数 $m<n$ 的多项式

$$a_0+a_1x+\cdots+a_mx^m$$

是 $p_0(x),\cdots,p_{n-1}(x)$的一个线性组合

$$a_0p_0(x)+a_1p_1(x)+\cdots+a_mp_m(x).$$

为了证明它们线性无关，考察线性关系

$$c_0p_0(x)+c_1p_1(x)+\cdots+c_{n-1}p^{n-1}(x)=\theta(x)$$

即

$$c_0+c_1x+\cdots+c_{n-1}x^{n-1}=0.$$

由我们多项式相等的定义，得出这样一个关系仅当 $c_0=c_1=\cdots=c_{n-1}=0$ 时才成立．这便建立了线性无关性而完成了证明．

为了理解上述证明，我们必须把这里涉及的多项式相等的问题与解关于 x 的一元方程的问题区别清楚．例如，可对 x 解方程

$$x^2-1=x+1,$$

而得到两解

$$x=2 \quad 和 \quad x=-1;$$

然而，多项式 x^2-1 不等于多项式 $x+1$．

例 19　证明 x 的所有多项式的向量空间 $\mathscr{P}$ 不是有限维空间．

设 $\mathscr{V}$ 是由有限个多项式生成的向量空间．那末显然 $\mathscr{V}$ 不能包含比每个生成元次数都高的多项式．这样 $\mathscr{V}$ 是 $\mathscr{P}$ 的一个真子空间，因此 $\mathscr{P}$ 不能是有限维的．

在下面的定理中，我们列举了一些关于基的有用的事实．

定理 II　设 $\mathscr{V}$ 是一个 n 维向量空间．

(a)　$\mathscr{V}$ 的任何 n 个线性无关向量构成 $\mathscr{V}$ 的一组基．

(b)　$\mathscr{V}$ 的任何 r 个线性无关向量是 $\mathscr{V}$ 一组基的一部分．

(c)　任何生成 $\mathscr{V}$ 的 n 个向量构成 $\mathscr{V}$ 的一组基．

证明　(a)假设 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 是 $\mathscr{V}$ 的线性无关向量．设 $\mathbf{w}$ 是 $\mathscr{V}$

的任何向量，则 $\mathbf{v}_1,\cdots,\mathbf{v}_n$，$\mathbf{w}$ 线性相关（例 16）。于是存在一个非平凡关系

$$c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n+d\mathbf{w}=\mathbf{0}.$$

在这关系里 $d\neq0$，因为如若不然，就存在一个 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 的非平凡线性关系，这与 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 的线性无关矛盾．因此

$$\mathbf{w}=-\left(\frac{c_1}{d}\right)\mathbf{v}_1-\cdots-\left(\frac{c_n}{d}\right)\mathbf{v}_n.$$

于是 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 生成 $\mathscr{V}$，并因此构成 $\mathscr{V}$ 的一组基。

(b)和(c)的证明留作练习。

例 20 证明多项式 $1+x,1+x^2$，$x+x^2$ 构成 $\mathscr{P}_3$ 的一组基．把 $2-3x+x^2$ 表达为这些基向量的线性组合。

由定理 II(a)，只要证明这些多项式线性无关即可．我们用三元数组(a_0,a_1,a_2)来表示 $\mathscr{P}_3$ 中的多项式

$$p(x)=a_0+a_1x+a_2x^2.$$

显然，多项式的加法和纯量乘法与相应的三元组看成 $\mathscr{R}^3$ 的向量时相应的运算一致．因此我们可以通过判定 $\mathscr{R}^3$ 中的相应相量 $(1,1,0)$，$(1,0,1)$和$(0,1,1)$的线性无关性，来检验 $1+x$，$1+x^2$，$x+x^2$ 的线性无关性．而 $\mathscr{R}^3$ 中上述三向量的线性无关性在 3.2 节例 12 中已验证．

最后，我们必须求出常数 c_1，c_2 和 c_3，使

$$2-3x+x^2=c_1(1+x)+c_2(1+x^2)+c_3(x+x^2).$$

由系数相等，我们得出

$$c_1+c_2=2,\quad c_1+c_3=-3,\quad c_2+c_3=1.$$

这个线性方程组有唯一解

$$c_1=-1,\quad c_2=3,\quad c_3=-2.$$

例 21 证明多项式

$$p_1(x)=2x+5x^2+x^3+\frac{1}{2}x^4,$$

$$p_2(x)=2+2x-\frac{1}{2}x^2+x^3+x^4,$$

$$p_3(x)=-8-x+12x^2+4x^3-3x^4,$$

$$p_4(x)=4+3x-x^2+2x^4$$

线性相关。

我们用 (a_0,a_1,a_2,a_3,a_4) 表示多项式 $a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$，则例 21 的四个多项式依次对应 3.2 节例 11 的四个向量。这就得出这些多项式适合非平凡线性关系

$$-2p_1(x)-2p_2(x)+p_3(x)+3p_4(x)=0$$

（见 3.2 节关系式(8)）。

练　　习

15. 下列哪些向量集合构成 $\mathscr{R}^3$ 的基？在构成基时，把向量(2,3,0)表作基向量的线性组合。

(i) (1,1,0),(2,2,0),(0,0,1)；

(ii) (1,−1,0), (1,0,−1), (0, 1, −1)；

(iii) (1,1,1), (1,−1,1), (1,−1,−1)；

(iv) (1,1,0),(1,0,0),(0,1,0),(0,0,1)。

[提示：见 3.2 节练习 8.]

16. 下列哪些多项式集合构成次数<3 的多项式向量空间的基？在构成基时，把 $2+3x$ 表作基向量的线性组合。

(i) $1+x, 2+2x, x^2$；

(ii) $1-x, 1-x^2, x-x^2$；

(iii) $1+x+x^2, 1-x+x^2, 1-x-x^2$；

(iv) $1+x, 1, x, x^2$。

[提示：见练习 15.]

17. 求例 21 四个多项式生成向量空间的维数。

18. 证明 3 次多项式集合不构成向量空间。

19. 验证向量(1,1,0,1),(1,0,1,1)和(1,1,1,0)线性无关。求出 $\mathscr{R}^4$ 的一个向量，使它和以上三个向量一起构成 $\mathscr{R}^4$ 的基。

20．证明定理 II(b)和(c)。

[提示：(b)[illegible] $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 线性无关．如果 $r=n$，则它们构成$\mathscr{V}$的一组基(定理 II(a))．$r>n$ 的情形是不可能的(例 16)．假设 $r<n$．由定理 I，$\mathbf{v}_1,\cdots,\mathbf{v}_r$ 不构成 $\mathscr{V}$ 的基．现在证明存在一个向量 $\mathbf{v}_{r+1}$ 使得 $\mathbf{v}_1,\cdots,\mathbf{v}_r,\mathbf{v}_{r+1}$ 线性无关．要证明(c)，用 3.2 节例 13 就推得和定理 I 矛盾．]

21．证明如果向量 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 构成向量空间$\mathscr{V}$的基，则$\mathscr{V}$中每个向量都可唯一地表示作 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 的线性组合．反之对吗？

22．用两种方法把 (2,−1,3) 表示为 (1,1,0)，(1,0,0)，(0,1,0) 和 (0,0,1)的线性组合．

3.4 无限维向量空间

我们在 3.3 节例 19 中发现，所有 x 的多项式的向量空间不能由有限个多项式生成．在物理和化学中大部分重要的向量空间都不是有限维的．为了讨论这些向量空间，我们必须把 3.1 节引入的一些概念一般化．

定义 I 设$\mathscr{V}$是向量空间，又设 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$是$\mathscr{V}$中向量的有限或无限集合．如果$\mathscr{V}$的每一个向量可表作向量 $\mathbf{v}_i(i=1,2,\cdots,n,\cdots)$ 的有限个数的线性组合，则我们称 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$ 生成$\mathscr{V}$．

如果$\mathscr{V}$不能由有限个向量生成，则我们称$\mathscr{V}$是无限维的．

例 22 所有实系数的 x 的多项式的空间 $\mathscr{P}$ 由多项式 $1,x,\cdots,x^n,\cdots$ 生成．$\mathscr{P}$是无限维的．

见 3.3 节例 18 和 19．

定义 II 如果 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$的一切有限集构成线性无关集，则称向量 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$是线性无关的．

例 23 $\mathscr{P}$中向量 $1,x,\cdots,x^n,\cdots$是线性无关的．

例 24 证明线性无关向量的一个无限集合生成一个无限维向量空间．

假设它们生成一个 n 维空间，则由 3.3 节例 16，这集合的任

意 $n+1$ 个向量都线性相关. 这是一个矛盾,于是证明完成.

定义 III 如果

(a) 向量 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$ 生成 $\mathscr{V}$, 且

(b) $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$ 线性无关,

则称向量空间 $\mathscr{V}$ 的向量 $\mathbf{v}_1,\cdots,\mathbf{v}_n,\cdots$ 构成 $\mathscr{V}$ 的一组基.

例 25 多项式 $1,x,\cdots,x^n,\cdots$ 构成 $\mathscr{P}$ 的一组基.

这立刻可由例 22 和 23 得出.

找出无限维空间基的问题不总是那么容易解决的. 例如考察在区间 $0\leqslant x\leqslant 1$ 内的一切连续函数的向量空间. 我们能不能找出这空间的一个基,即我们能不能找出一个线性无关函数集合,使得在 $0\leqslant x\leqslant 1$ 连续的任何函数可表作这集合中有限个函数的线性组合呢? 如果你考虑到这空间里连续函数的多样性 (从 $x=0$ 到 $x=1$ 的任何连续曲线确定一个连续函数), 你可能认识到这是一个困难的问题. 实际上, 要解答这个问题就超出本书的范围了.

练　习

23. 下列哪些多项式无限集构成 $\mathscr{P}$ 的一组基?

(i) $1,1+x,1+x+x^2,1+x+x^2+x^3,\cdots$.

(ii) $1+x,x+x^2,x^2+x^3,x^3+x^4,\cdots$.

(iii) $1,1+x,x+x^2,x^2+x^3,x^3+x^4,\cdots$.

在一个集合构成基的时候, 把多项式 $1+2x^3-x^5$ 表作有限个基向量的线性组合.

24. 证明三角函数 $\sin x,\sin 2x,\sin 3x,\cdots$ 是线性无关的.

[提示: 假设对某个正整数 n 有线性关系

$$c_1\sin x+c_2\sin 2x+\cdots+c_n\sin nx=0,$$

对一切 x 成立. 由计算

$$\int_{-\pi}^{\pi}(c_1\sin x+c_2\sin 2x+\cdots+c_r\sin rx+\cdots+c_n\sin nx)\sin rx\,dx,$$

证明对一切 $r(1\leqslant r\leqslant n)$, $c_r=0$.]

第四章　用向量方法的坐标几何

在本章中，我们将用向量方法研究三维坐标几何．我们引进两种向量“乘法”：

（i）点积或纯量积；及

（ii）叉积或向量积．

两个向量的点积是一个纯量，此纯量与这两个向量的长度及它们的夹角有关．点积的概念是很重要的，这是因为它不仅在坐标几何里有用，而且还出现在物理问题中（例如一个力所作的功是一个点积），还因为高维空间中的正交性这个重要概念正是通过推广点积来讨论的．下一章里，我们要讨论正交性．

两个向量的叉积是一个向量．象点积一样，在物理里它有重要的应用（例如力矩是一个叉积）．然而，和点积不同的是，叉积不能直接推广到异于三维的空间上去．

除非另有说明，本章的所有向量都假定是三维的．它们由一个长度和一个方向来确定，或等价地由一个三元数组 $\mathbf{v}=(v_1, v_2, v_3)$ 来确定，此三元数组给出了这个向量在某直角坐标系中的三个分量．

4.1　三维空间中的向量的长度和方向

为了表示三维空间中的点，我们引入一个右手直角坐标系 $O1$，$O2$，$O3$（图 18）．坐标系之所以称为右手系，是因为如果使右手的姆指和食指依次指向 $O1$ 和 $O2$ 方向，那末右手的中指可指向 $O3$ 方向．用左手是不能做到这一点的！

设 P 点的位置向量是 $\mathbf{v}$．关于这个坐标系，(v_1, v_2, v_3) 能表示 $\mathbf{v}$ 和 P 二者．数 v_1, v_2, v_3 称为这个坐标系下 P 的坐标和 $\mathbf{v}$ 的分

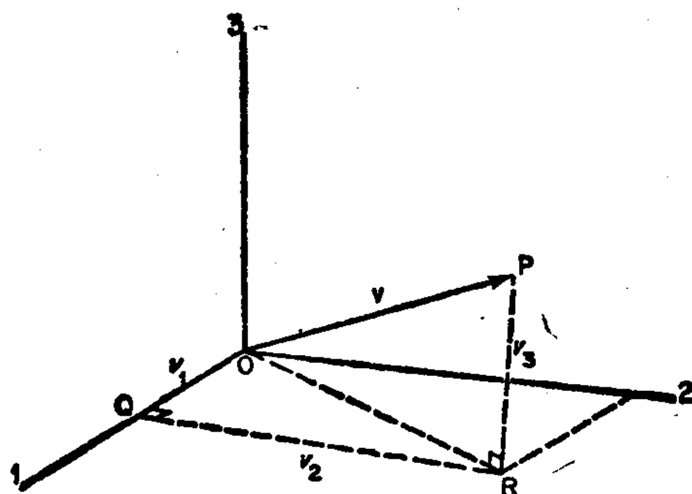


图 18 右手直角坐标系

量.

由

$$\|\mathbf{v}\|=\sqrt{(v_1^2+v_2^2+v_3^2)} \tag{1}$$

给出 v 的长度.

证明 在图 18 里,角 ORP 和 OQR 是直角. 因此,由毕达哥拉斯定理

$$\|\mathbf{v}\|^2=OP^2=OR^2+RP^2=OQ^2+QR^2+RP^2=v_1^2+v_2^2+v_3^2,$$

这就是所要证明的.

注意 $\|k\mathbf{v}\|=|k|\,\|\mathbf{v}\|$,对任意纯量 k 成立.

有单位长度的向量 u 称为单位向量. 向量

$$\mathbf{e}_1=(1,0,0), \quad \mathbf{e}_2=(0,1,0), \quad \mathbf{e}_3=(0,0,1)$$

是坐标轴方向的单位向量. 常常依次用 i, j, k 分别表示它们. 向量$\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$也是单位向量.

设 v 是任意非零向量. 向量

$$\mathbf{u}=\frac{1}{\|\mathbf{v}\|}\mathbf{v} \tag{2}$$

是 v 方向上的单位向量.

作出向量 **v** 方向上的单位向量的过程称为正规化 **v**.

方向余弦

设 $\mathbf{v}=(v_1,v_2,v_3)$ 是一个任意向量，在图 19 中用 $\overrightarrow{OP}$ 表示.

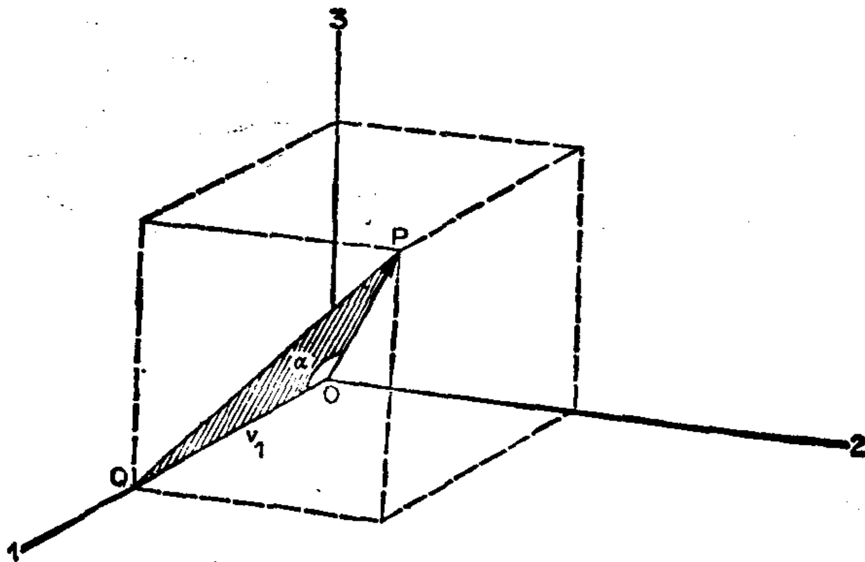


图 19　方向余弦

设 α 是 OP 和 $O1$ 轴的夹角. 因为 QP 位于一个与包含 $O2$ 和 $O3$ 轴的平面相平行的平面中，所以角 OQP 是直角.

因此

$$\cos\alpha=\frac{OQ}{OP}=\frac{v_1}{\|\mathbf{v}\|}.$$

类似地，如果 β 和 γ 依次表示 OP 和 $O2$，OP 和 $O3$ 的夹角，则

$$\cos\beta=\frac{v_2}{\|\mathbf{v}\|}\quad\text{和}\quad\cos\gamma=\frac{v_3}{\|\mathbf{v}\|}.$$

$\cos\alpha$，$\cos\beta$ 和 $\cos\gamma$ 称为向量

$$\mathbf{v}=(v_1,v_2,v_3)$$

的方向余弦.

通常用字母 l，m 和 n 表示方向余弦. 立即可推出下列关于

方向余弦的一些事实.

（i）两个向量 v 和 w 有相同的方向余弦当且仅当 w 是 v 的正倍数.

（ii）如果 u 是正规化 v 所得的单位向量，即

$$\mathbf{u}=\frac{1}{\|\mathbf{v}\|}\mathbf{v},$$

则 $\mathbf{u}=(l, m, n)$，这里 l, m, n 是 v 的（也是 u 的）方向余弦 $\cos\alpha, \cos\beta, \cos\gamma$.

(iii) $(\cos\alpha)^2+(\cos\beta)^2+(\cos\gamma)^2=l^2+m^2+n^2=1$. (3)

例 1 求 $\mathbf{v}=(2,-2,1)$ 的方向余弦.

正规化 v，得到和 v 有相同方向的单位向量 $\frac{1}{3}(2,-2,1)$. v 的方向余弦是

$$\cos\alpha=\frac{2}{3}, \quad \cos\beta=-\frac{2}{3}, \quad \cos\gamma=\frac{1}{3}.$$

α 和 γ 是锐角，而 β 是钝角.

注意 $-\mathbf{v}=(-2,2,-1)$ 的方向余弦是

$$\cos\alpha'=-\frac{2}{3}, \quad \cos\beta'=+\frac{2}{3}, \quad \cos\gamma'=-\frac{1}{3},$$

这和关系 $\alpha'=\pi-\alpha, \beta'=\pi-\beta, \gamma'=\pi-\gamma$ 相符合(图 20).

例 2 两点间的距离

P 和 Q 的位置向量依次是 v 和 w（图 21）. 求 P 和 Q 之间的距离.

$\overrightarrow{PQ}=\mathbf{w}-\mathbf{v}$. 因此 P 和 Q 之间的距离是 $\|\mathbf{w}-\mathbf{v}\|$.

用坐标的语言来说，如果 P 和 Q 依次是点 (v_1, v_2, v_3) 和 (w_1, w_2, w_3)，则

$$PQ=\sqrt{(w_1-v_1)^2+(w_2-v_2)^2+(w_3-v_3)^2}.$$

长度 $\|\mathbf{w}-\mathbf{v}\|$ 又称为向量 v 和 w 之间的距离.

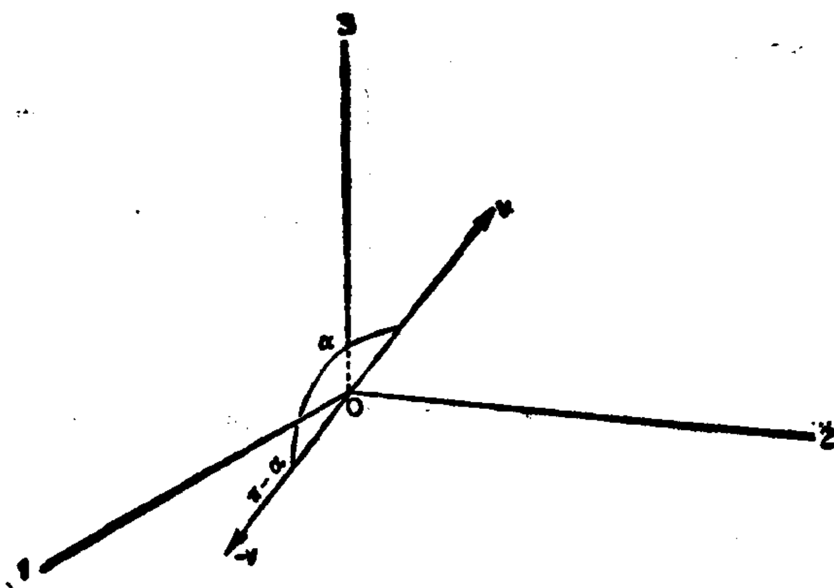


图 20

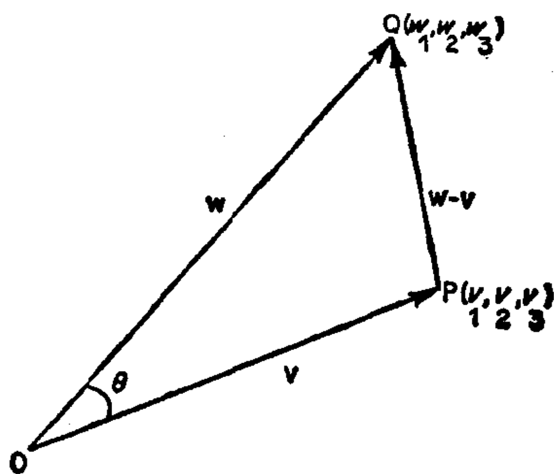


图 21 P和Q的距离

练　　习

1．求向量 $\mathbf{v}=(1,-1,2)$ 的长度和方向余弦．正规化 $\mathbf{v}$ 得到什么向量？

2．对向量 $-\mathbf{v}=(-1,1,-2)$ 和 $3\mathbf{v}=(3,-3,6)$ 重复练习 1．

3．证明四点 $P(1,1,1)$，$Q(2,0,3)$，$R(0,1,2)$，$S(-1,2,0)$ 构成一个平行四边形 $PQRS$．求对角线 PR 和 QS 的长度．

4.2 点积(或纯量积)

设 $\mathbf{v}=(v_1,v_2,v_3)$ 和 $\mathbf{w}=(w_1,w_2,w_3)$ 依次是点 P 和 Q 对于一个直角坐标系的位置向量(见图 21). 假设 $\mathbf{v}\neq\mathbf{0}$ 和 $\mathbf{w}\neq\mathbf{0}$,因此 P 和 Q 不同于原点 O. 我们定义向量 $\mathbf{v}$ 和 $\mathbf{w}$ 的夹角为角 $POQ(=\theta)$,使 $0\leqslant\theta\leqslant\pi$.

由余弦定理

$$PQ^2=OP^2+OQ^2-2\,OP\cdot OQ\cos\theta. \tag{4}$$

我们有

$$PQ^2=\|\mathbf{w}-\mathbf{v}\|^2=(w_1-v_1)^2+(w_2-v_2)^2+(w_3-v_3)^2,$$
$$OP^2=\|\mathbf{v}\|^2=v_1^2+v_2^2+v_3^2$$

和

$$OQ^2=\|\mathbf{w}\|^2=w_1^2+w_2^2+w_3^2.$$

代到等式(4)里并化简,我们得到

$$\cos\theta=\frac{v_1w_1+v_2w_2+v_3w_3}{\|\mathbf{v}\|\|\mathbf{w}\|} \tag{5}$$

这个表达式的分子称为向量 $\mathbf{v}=(v_1,v_2,v_3)$ 和 $\mathbf{w}=(w_1,w_2,w_3)$ 的点积或纯量积,记作

$$\mathbf{v}\cdot\mathbf{w}=v_1w_1+v_2w_2+v_3w_3. \tag{6}$$

此式也定义了当 $\mathbf{v}=\mathbf{0}$ 或 $\mathbf{w}=\mathbf{0}$ 时 $\mathbf{v}$ 和 $\mathbf{w}$ 的点积. 于是,对一切向量 $\mathbf{v}$ 都有 $\mathbf{0}\cdot\mathbf{v}=\mathbf{v}\cdot\mathbf{0}=0$.

可把公式(5) 表达为

$$\mathbf{v}\cdot\mathbf{w}=\|\mathbf{v}\|\|\mathbf{w}\|\cos\theta. \tag{7}$$

常把这个式子作为两个非零向量 $\mathbf{v}$ 和 $\mathbf{w}$ 点积的定义. 在这个定义中不需要参照一个特殊的坐标系.

例 3 求向量 $\mathbf{v}=(1,-2,2)$ 和 $\mathbf{w}=(0,3,4)$ 的夹角 θ.

$$\|\mathbf{v}\|=\sqrt{(1+4+4)}=3,\quad \|\mathbf{w}\|=5,$$

和

$$\mathbf{v}\cdot\mathbf{w}=(1\cdot 0)+(-2\cdot 3)+(2\cdot 4)=2.$$

代入公式(7)，我们得到 $\cos\theta=\frac{2}{15}$，由此可求出 θ.

正交性

我们称两个非零向量是正交的(或垂直的)当且仅当它们之间的夹角是一个直角. 由公式(7)，这等价于它们的点积等于零.

说零向量 $\mathbf{0}$ 和另一个向量构成一个角没有什么意义，因为我们不能给 $\mathbf{0}$ 分配一个方向. 但我们将采用 $\mathbf{0}$ 正交于每个向量的约定. 这样，我们可无例外地叙述：

定义 I 两个向量 $\mathbf{v}$ 和 $\mathbf{w}$ 是正交的，当且仅当 $\mathbf{v}\cdot\mathbf{w}=0$.

例 4 验证 $\mathbf{v}=(1,2,1)$ 正交于 $\mathbf{w}=(2,-2,2)$.

$$\mathbf{v}\cdot\mathbf{w}=1\cdot 2+2\cdot(-2)+1\cdot 2=0.$$

根据我们正交性的定义，向量 $\mathbf{0}$ 和自身正交. 没有别的向量有这个性质. 因为若

$$\mathbf{v}=(v_1,v_2,v_3)\neq\mathbf{0},$$

则

$$\mathbf{v}\cdot\mathbf{v}=v_1^2+v_2^2+v_3^2>0.$$

我们已经顺便证明了可用点积表示出长度：

$$\|\mathbf{v}\|^2=\mathbf{v}\cdot\mathbf{v}. \tag{8}$$

例 5 假设 $\mathbf{a},\mathbf{b}$ 和 $\mathbf{c}$ 是适合 $\mathbf{a}\cdot\mathbf{c}=\mathbf{b}\cdot\mathbf{c}$ 的向量. 这可推出 $\mathbf{a}=\mathbf{b}$ 吗？

不！我们不能约去 $\mathbf{c}$. 例如若

$$\mathbf{e}_1=(1,0,0),\quad \mathbf{e}_2=(0,1,0)\quad 和\quad \mathbf{e}_3=(0,0,1),$$

则

$$\mathbf{e}_1\cdot\mathbf{e}_3=\mathbf{e}_2\cdot\mathbf{e}_3=0,$$

但 $\mathbf{e}_1\neq\mathbf{e}_2$.

在以下定理中，我们列举了点积的三个代数性质. 它们给出了运算规律，同时也是我们将在下一章考察的一般情形的基础.

定理 I 点积有以下性质：

(a) 它是对称的：$\mathbf{v}\cdot\mathbf{w}=\mathbf{w}\cdot\mathbf{v}$；

(b) 它是双线性的：$\mathbf{s}\cdot(k\mathbf{v}+l\mathbf{w})=k(\mathbf{s}\cdot\mathbf{v})+l(\mathbf{s}\cdot\mathbf{w})$，

$$(k\mathbf{v}+l\mathbf{w})\cdot\mathbf{s}=k(\mathbf{v}\cdot\mathbf{s})+l(\mathbf{w}\cdot\mathbf{s})；$$

(c) 它是正定的：若 $\mathbf{v}\neq\mathbf{0}$，则 $\mathbf{v}\cdot\mathbf{v}>0$；且 $\mathbf{0}\cdot\mathbf{0}=0$。

从点积的定义(6)容易验证这些性质，注意性质(c)仅仅是叙述了一切非零向量有正长度.

例 6 证明：如果 $\mathbf{a}\cdot\mathbf{c}=\mathbf{b}\cdot\mathbf{c}$ 对一切向量 $\mathbf{c}$ 成立，则 $\mathbf{a}=\mathbf{b}$（和例5相比较）.

从 $\mathbf{a}\cdot\mathbf{c}=\mathbf{b}\cdot\mathbf{c}$ 我们得到 $(\mathbf{a}-\mathbf{b})\cdot\mathbf{c}=0$（用定理 I (b)）。于是 $\mathbf{a}-\mathbf{b}$ 正交于 $\mathbf{c}$，对一切向量 $\mathbf{c}$ 是正确的. 特别，$\mathbf{a}-\mathbf{b}$ 正交于自身，而除了 $\mathbf{0}$，没有其它向量具有这个性质. 因此 $\mathbf{a}=\mathbf{b}$。

练　习

4．求夹角：

(i) 向量(1,2,2)和(2,−1,−2)；

(ii) 向量(1,2,2)和(−2,1,2)。

并加以比较。

5．证明三点 $P(1,1,-1)$，$Q(3,2,1)$，$R(-1,-1,0)$ 构成一个等腰三角形 PQR。求出角 QPR。再求出 QR 中点 M 的坐标，并验证 PM 垂直于 QR。

[提示：用 $\mathbf{p},\mathbf{q},\mathbf{r}$ 表示 P,Q,R 的位置向量。则 $\overrightarrow{PQ}=\mathbf{q}-\mathbf{p}$，等等.]

6．求一个向量，使它正交于 $\mathbf{v}=(1,2,2)$ 和 $\mathbf{w}=(-2,-1,2)$ 两者. 证明恰有两个单位向量正交于 $\mathbf{v}$ 和 $\mathbf{w}$ 两者。

7．设 $\mathbf{e}_1=(1,0,0)$，$\mathbf{e}_2=(0,1,0)$，$\mathbf{e}_3=(0,0,1)$。（这些是坐标轴方向的单位向量.）验证当 $i\neq j$ 时，$\mathbf{e}_i\cdot\mathbf{e}_j=0$，和 $\mathbf{e}_i\cdot\mathbf{e}_i=1$。

可把任何向量 $\mathbf{v}$ 表作

$$\mathbf{v}=(v_1,v_2,v_3)=v_1\mathbf{e}_1+v_2\mathbf{e}_2+v_3\mathbf{e}_3.$$

验证

$$v_i=\mathbf{v}\cdot\mathbf{e}_i\quad(i=1,2,3).$$

8. 设向量 $\mathbf{b}_1,\mathbf{b}_2$ 和 $\mathbf{b}_3$ 互相正交且是非零的. 证明它们是线性无关的.

[提示：假设一个线性关系

$$c_1\mathbf{b}_1+c_2\mathbf{b}_2+c_3\mathbf{b}_3=\mathbf{0}.$$

取和 $\mathbf{b}_1$ 的点积就证明了 $c_1=0$,等等.]

9. 已知非零向量 $\mathbf{a}$ 和 $\mathbf{b}$,求形如 $\mathbf{a}-k\mathbf{b}$ 的最短向量.

[提示：在图上试对各种值 k 画 $\mathbf{a}-k\mathbf{b}$. 求得 $\mathbf{a}-k\mathbf{b}$ 尽可能短的简单几何准则.]

10. 求顶点为 $P(1,1,1)$, $Q(2,0,3)$, $R(0,1,2)$, $S(-1,2,0)$ 的平行四边形 $PQRS$ 对角线 PR 和 QS 的夹角.

11. 证明 $\mathbf{a}-\mathbf{b}$ 和 $\mathbf{a}+\mathbf{b}$ 正交的充分必要条件是 $\|\mathbf{a}\|=\|\mathbf{b}\|$.

[提示：用定理 I 和公式(8).]

几何中相应定理是什么？试加以说明.

12. 证明 $\mathbf{a}$ 和 $\mathbf{b}$ 正交的充分必要条件是

$$\|\mathbf{a}+\mathbf{b}\|^2=\|\mathbf{a}\|^2+\|\mathbf{b}\|^2.$$

几何中相应定理是什么？

13. 证明任意平行四边形中对角线平方和等于相邻两边平方和的两倍.

14. 从 A 移动到 B，常力 $\mathbf{F}$ 所作的功(图 22)等于 $\mathbf{F}$ 在运动方向分量数值和运动距离的乘积. 把它表示为点积.

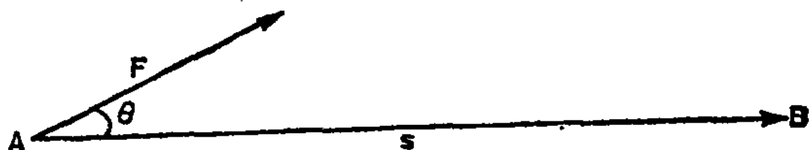


图 22 一个力所做的功

15. 证明，若定义 $\mathbf{v}=(v_1,v_2)$ 和 $\mathbf{w}=(w_1,w_2)$ 点积为

$$\mathbf{v}\cdot\mathbf{w}=v_1w_1+v_2w_2,$$

公式(7)和(8)对二维向量有效.

16. 本练习涉及在一个直角坐标系中表达的二维向量.

(i) 求 $\mathbf{v}=(3,4)$ 和 $\mathbf{w}=(5,-12)$ 的夹角.

(ii) 求正交于 $\mathbf{v}=(3,4)$ 的一个向量. 证明恰有两个单位向量正交于 $\mathbf{v}$.

用图说明你的解.

4.3 三维空间中的平面方程

从原点到平面所画垂线长 $p=OP$ 和$\overrightarrow{OP}$ 方向的单位向量 $\mathbf{u}=(l,m,n)$决定了三维空间中的一个平面(图 23). 我们把这个向量称为平面的单位法向量. 如果平面通过原点($p=0$), 我们选垂直于平面的两个单位向量中的任何一个(记作 $\mathbf{u}$) 作为单位法向量(没取的另一个单位法向量是$-\mathbf{u}$).

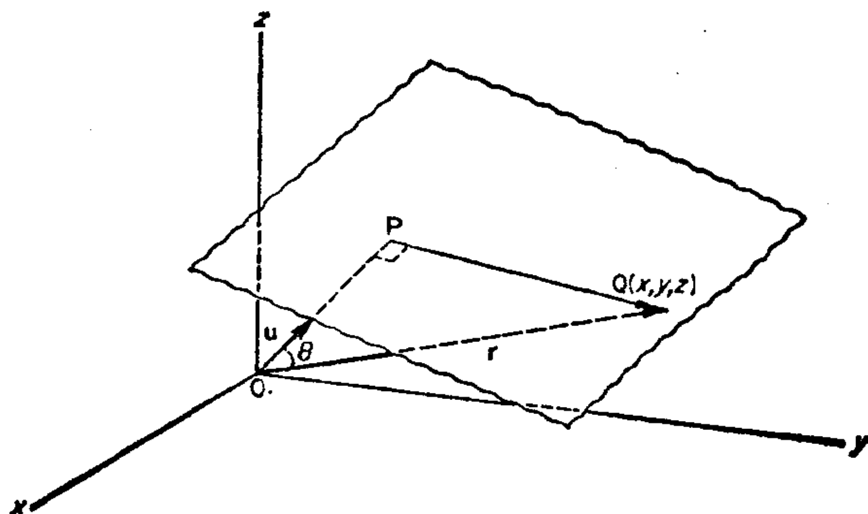


图 23 平面方程

一点* $Q(x,y,z)$在平面中当且仅当 $\overrightarrow{PQ}$ 正交于 $\mathbf{u}$. 设 $\mathbf{r}=\overrightarrow{OQ}$. 于是 $\overrightarrow{PQ}=\overrightarrow{OQ}-\overrightarrow{OP}=\mathbf{r}-p\mathbf{u}$. 因此这平面的方程是

$$\mathbf{u}\cdot(\mathbf{r}-p\mathbf{u})=0. \tag{9}$$

由此

$$\mathbf{u}\cdot\mathbf{r}=p(\mathbf{u}\cdot\mathbf{u})=p$$

(记住 $\mathbf{u}$ 是单位向量). 把 $\mathbf{u}=(l,m,n)$和

$$\mathbf{r}=(x,y,z)$$

代入,我们得到这平面方程形如

---

* 本章到这里,我们已经用了容易推广到 n 维空间的记号. 现在我们改成三维几何更通用的 x, y, z 记号.

$$lx+my+nz=p, \tag{10}$$

这里 p 是从 O 到这平面的垂直距离，$\mathbf{u}=(l,m,n)$ 是这平面的单位法向量.

当 $p=0$，平面方程是 $lx+my+nz=0$。如果在这情形我们取 $-\mathbf{u}=(-l,-m,-n)$ 作为单位法向量，我们得到方程 $-lx-my-nz=0$，它显然确定同一个平面.

注意，当 a，b，c 不全为零时，x，y，z 的任意线性方程

$$ax+by+cz=d \tag{11}$$

代表了一个平面.

证明 我们可以假设 $d\geqslant 0$，因为如果 $d<0$，我们可用 -1 乘方程(11)每项. 用非零数

$$k=\sqrt{(a^2+b^2+c^2)}$$

除(11)我们得到

$$lx+my+nz=p, \tag{12}$$

这里 $l=\frac{a}{k}$，$m=\frac{b}{k}$，$n=\frac{c}{k}$，$p=\frac{d}{k}\geqslant 0$ 且 $l^2+m^2+n^2=1$.我们看出方程(12)(因而方程(11))是这样一个平面方程：这平面和原点的垂直距离是 p，单位法向量是 $\mathbf{u}=(l,m,n)$，我们把方程(12)称为方程(11)的标准形式. 注意向量(a,b,c)垂直于方程(11)所确定的平面.

例 7 描述由下列方程确定的平面：

(i) $2x+3y-4z=5$；及

(ii) $2x+3y-4z=-5$.

(i) 的标准形式是

$$\frac{2}{\sqrt{29}}x+\frac{3}{\sqrt{29}}y-\frac{4}{\sqrt{29}}z=\frac{5}{\sqrt{29}},$$

而(ii)的标准形式是

$$-\frac{2}{\sqrt{29}}x-\frac{3}{\sqrt{29}}y+\frac{4}{\sqrt{29}}z=-\frac{5}{\sqrt{29}}.$$

这两个平面是互相平行的. 每一个平面都在离原点距离为 $\frac{5}{\sqrt{29}}$ 处. 从 O 画出的单位向量 $\mathbf{u}=\left(\frac{2}{\sqrt{29}},\frac{3}{\sqrt{29}},-\frac{4}{\sqrt{29}}\right)$ 指向平面(i). 单位向量 $-\mathbf{u}$ 指向平面(ii). 两平面间的距离是 $\frac{10}{\sqrt{29}}$.

例 8 描述方程为

$$3x+4y=5$$

的平面.

这个平面包含 xy 平面中的直线 L: $3x+4y=5$, 且平行于 z-轴(图 24).

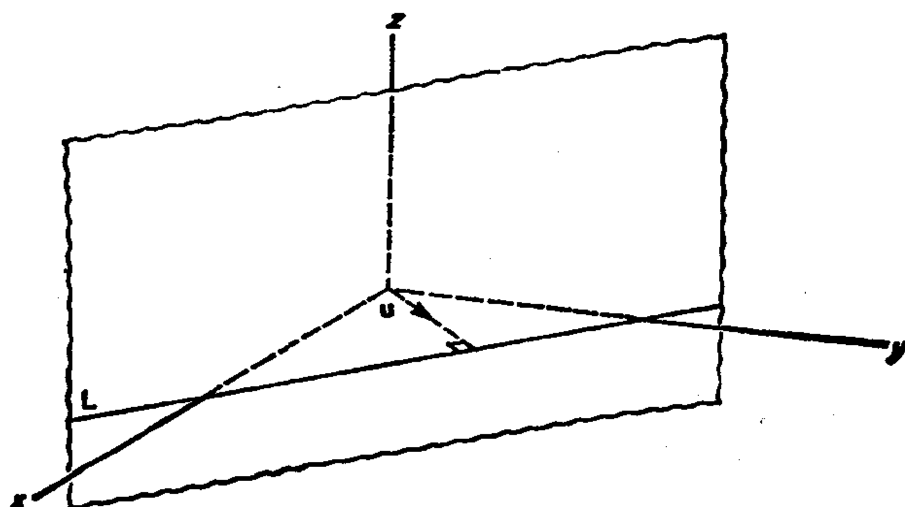


图 24 平面 $3x+4y=5$

这平面的标准形式方程是

$$\frac{3}{5}x+\frac{4}{5}y+0z=1.$$

O 到这平面的垂直距离是 1. 这平面的单位法向量是

$$\mathbf{u}=\left(\frac{3}{5},\frac{4}{5},0\right).$$

平面波

当通过介质传播一个扰动的时候，波动便发生了. 熟悉的例子是丢一颗小石子到水池里引起的水波，在空气中传播的声波，以及在以太中传播的电磁波（无线电波，光波）. 在一个直角坐标系中，要描述一个波动，只要说明每一点(x,y,z)在每一时刻t的扰动就可以了. 这个扰动记为$U(x,y,z,t)$.

以常速v沿着单位向量$\mathbf{u}$方向传播的平面波具有下列性质：

(i) 给了任何垂直于$\mathbf{u}$的平面Π和任何固定时刻，则Π的一切点上扰动是相同的.

(ii) 经过时间t以后扰动已在$\mathbf{u}$方向传播了距离vt.

假设$\mathbf{u}=(l,m,n)$. 于是Π的方程是

$$lx+my+nz=q,$$

这里q是某个常数(见图 25). (注意在这个图示中q是负的，因为$\mathbf{u}$指离Π).

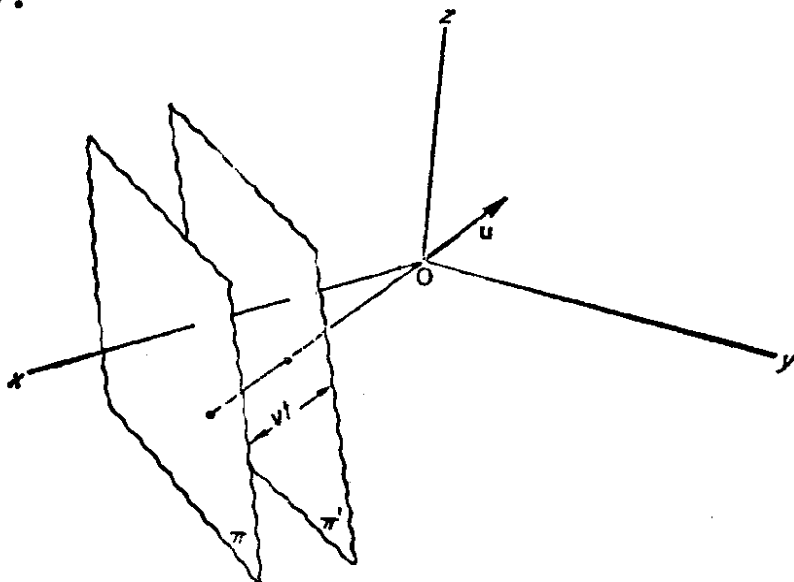


图 25 平面波

经过时间t后，在平面Π'上将观察到原来在Π上的扰动，Π'的方程是

$$lx+my+nz=q+vt.$$

现在设 f 是一个任意的一元函数．于是

$$U(x,y,z,t)=f(lx+my+nz-vt) \tag{13}$$

表示了如上描述的一个平面波动，因为原先在平面 Π 上用 U 来度量的扰动和经过时间 t 后在 Π' 上度量的扰动是相同的；两者都是 $f(q)$．

在 1.3 节例 4 中，我们引进过波动方程

$$\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}-\frac{1}{v^2}\frac{\partial^2 U}{\partial t^2}=0 . \tag{14}$$

现在我们证明平面波函数(13)适合这波动方程，由此说明给方程(14)取这样的名字是有理由的．

证明 对 x 求偏导数我们得到

$$\frac{\partial U}{\partial x}=lf'(lx+my+nz-vt)$$

和

$$\frac{\partial^2 U}{\partial x^2}=l^2 f''(lx+my+nz-vt),$$

这 f' 和 f'' 各表示函数 f 的一阶和二阶导数．(例如，如果

$$U=\sin(lx+my+nz-vt),\ \text{则}$$

$$\frac{\partial U}{\partial x}=l\cdot\cos(lx+my+nz-vt),\ \text{等等.})$$

类似地我们计算 $\frac{\partial^2 U}{\partial y^2}$，$\frac{\partial^2 U}{\partial z^2}$ 和 $\frac{\partial^2 U}{\partial t^2}$，得到

$$\begin{aligned}&\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}-\frac{1}{v^2}\frac{\partial^2 U}{\partial t^2}\\&\quad=\left[(l^2+m^2+n^2)-\frac{v^2}{v^2}\right]f''(lx+my+nz-vt)\\&\quad=0,\end{aligned}$$

因为 $l^2+m^2+n^2=1$．

例 9 描述平面波

$$U=\sin(x\cos\alpha+y\sin\alpha-vt). \tag{15}$$

$\mathbf{u}=(\cos\alpha,\sin\alpha,0)$ 给出了波传播的方向. 相等扰动的一系列平面 Π 平行于 z-轴. 为了方便, 可画一草图, 把 x-轴与 y-轴画在纸平面上(图 26).

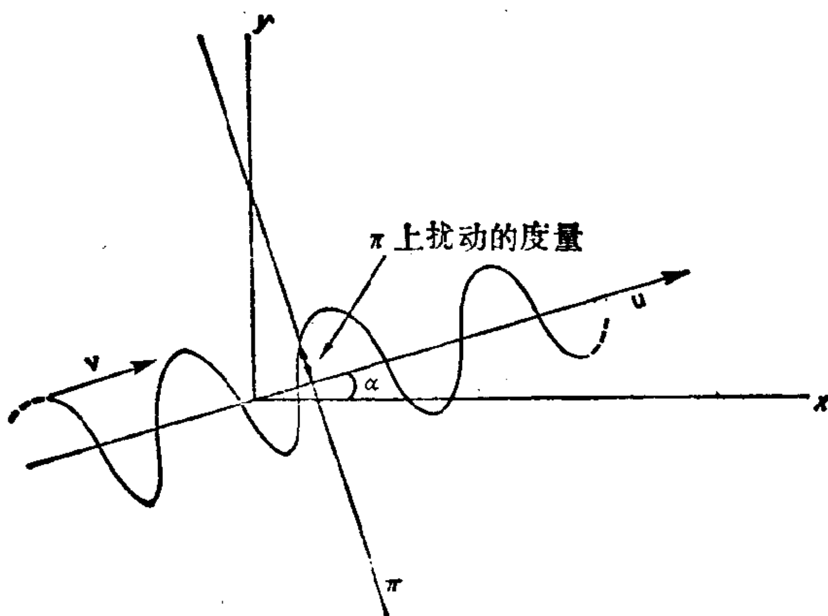


图 26 平面波 $U=\sin(x\cos\alpha+y\sin\alpha-vt)$

可以把方程(15)当作两维波(例如, 一个水池表面的波). 这种情形下图 26 中 Π 表示为一条等扰动的直线. 波动的速度是 v.

如果 $\alpha=0$, 方程(15)表示了由

$$U=\sin(x-vt)$$

确定的沿 x-轴的正弦波运动.

练　　习

17. 描述方程

$$x+2y-2z=1 \quad 和 \quad 2x+y+2z=3$$

确定的平面.

证明它们相交成直角.

[提示: 考察它们的单位法向量.]

18. 证明平面

$$x+2y-2z=3 \quad 和 \quad x+2y-2z=-3$$

是平行的，求出它们之间的垂直距离.

19．求含点 $A(1,2,-2)$ 且平行于平面

$$x+2y+2z=3$$

的平面方程.

[提示：这方程一定形如 $x+2y+2z=k$. 为什么?]

20．求点 $A(1,2,-2)$ 到平面

$$x+2y+2z=3$$

的垂直距离.

[提示：应用练习 19 和练习 18 的方法.]

21．一个平面通过 $(8,-2,2)$，$(2,1,-4)$ 和 $(2,4,-6)$. 求它的方程. 求原点到它的垂直距离和单位法向量的方向余弦.

22．证明可以把平面波即函数(13)表示为

$$U=f(\mathbf{u}\cdot\mathbf{r}-vt)$$

的形式.

23．证明

$$U=a\sin(x-2y+2z-6vt)$$

确定一个在单位向量 $\mathbf{u}=\left(\frac{1}{3},-\frac{2}{3},\frac{2}{3}\right)$ 方向上以 $2v$ 速度运动的平面波. 再证明扰动波长是 $\frac{2\pi}{3}$，振幅是 a.

[提示：证明可把波动表示为

$$U=a\sin 3(\mathbf{u}\cdot\mathbf{r}-2vt).$$

这是(13)形式的函数，这里 $f(q)=a\sin 3q$ 且用 $2v$ 代替了 v. 要确定波长，只要固定 t，并考察沿着传播轴 $\mathbf{r}=k\mathbf{u}$ 两个相似扰动点之间的距离.]

24．描述平面波

(i) $U=a\sin(3x-4y-vt)$；

(ii) $U=a\sin(3x-4y+vt)$；

(iii) $U=a\sin(4x+3y-vt)$.

并加以比较.

要了解波动的简单计算，可见 C. A. Coulson, *Waves: a Mathematical Account of the Common Types of Wave*

Motion, Seventh Edition, Oliver & Boyd, 1955.

4.4 向量积(或叉积)

问题 $\mathbf{v}=(a,b,c)$和$\mathbf{v}'=(a',b',c')$是由一个右手直角坐标系所表示的两个非零不平行的向量. 求与$\mathbf{v}$, $\mathbf{v}'$都垂直的一个向量.

所求向量比如说是$\mathbf{r}=(x, y, z)$, 必须适合$\mathbf{v}\cdot\mathbf{r}=0$和$\mathbf{v}'\cdot\mathbf{r}=0$,即

$$ax+by+cz=0$$

和

$$a'x+b'y+c'z=0 \tag{16}$$

用c'乘第一个方程,用c乘第二个方程,相减消去z. 得

$$(bc'-cb')y=(ca'-ac')x.$$

类似地, 消去y我们得到

$$(ab'-ba')x=(bc'-cb')z.$$

因此显然,对任何纯量k,

$$x=k(bc'-cb'),\quad y=k(ca'-ac'),\quad z=k(ab'-ba')$$

即

$$\mathbf{r}=k(bc'-cb',ca'-ac',ab'-ba') \tag{17}$$

是方程组(16)的解. 这实际上是(16)的通解, 因为它显然包含垂直于$\mathbf{v}$和$\mathbf{v}'$这一(唯一)方向的所有向量.

在方程(17)中取$k=1$, 所得到的向量称为$\mathbf{v}$和$\mathbf{v}'$的向量积或叉积,并记作*$\mathbf{v}\times\mathbf{v}'$.

于是

$$\mathbf{v}\times\mathbf{v}'=(bc'-cb',ca'-ac',ab'-ba'). \tag{18}$$

即使$\mathbf{v}$和$\mathbf{v}'$平行或等于$\mathbf{0}$,我们仍用公式(18)定义$\mathbf{v}$和$\mathbf{v}'$的向量

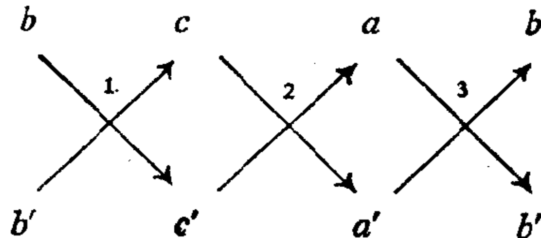
* 另一个通用的符号是$\mathbf{v}\wedge\mathbf{v}'$.

积. 从公式(18)立刻得出

(a) $\mathbf{v}\times\mathbf{0}=\mathbf{0}\times\mathbf{v}=\mathbf{0}$,对一切 $\mathbf{v}$ 成立;

(b) 如果 $\mathbf{v}$ 和 $\mathbf{v}'$ 是适合 $\mathbf{v}'=k\mathbf{v}$ 的非零向量(即 $\mathbf{v}$ 和 $\mathbf{v}'$ 是平行的),则 $\mathbf{v}\times\mathbf{v}'=\mathbf{0}$.

下图可以帮助你记住公式(18).

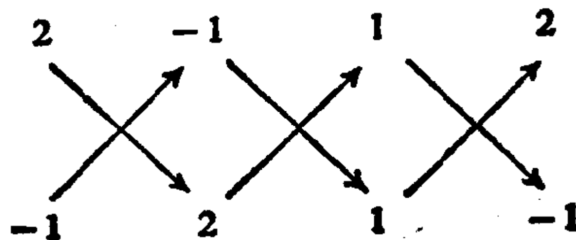


$$(a,b,c)\times(a',b',c')=(bc'-cb',ca'-ac',ab'-ba').$$

向下的箭头联系着正的积,向上箭头联系着负的积. 1, 2, 3 标明了分量数.

例 10 求一个向量垂直于两个向量

$$\mathbf{v}=(1,2,-1) \quad 和 \quad \mathbf{v}'=(1,-1,2).$$



由图我们计算

$$\mathbf{v}\times\mathbf{v}'=(4-1,-1-2,-1-2)=(3,-3,-3).$$

这个向量(及它的每个纯量倍)垂直于 $\mathbf{v}$ 和 $\mathbf{v}'$. 可以很快验证下述结果: $\mathbf{v}\cdot(\mathbf{v}\times\mathbf{v}')=0$ 及 $\mathbf{v}'\cdot(\mathbf{v}\times\mathbf{v}')=0$.

例 11 $\mathbf{i},\mathbf{j},\mathbf{k}$ 是右手直角坐标系坐标轴 Ox,Oy,Oz 方向的单位向量(图 27). 验证

$$\mathbf{i}\times\mathbf{j}=\mathbf{k}=-\mathbf{j}\times\mathbf{i},$$

$$\mathbf{j}\times\mathbf{k}=\mathbf{i}=-\mathbf{k}\times\mathbf{j},$$

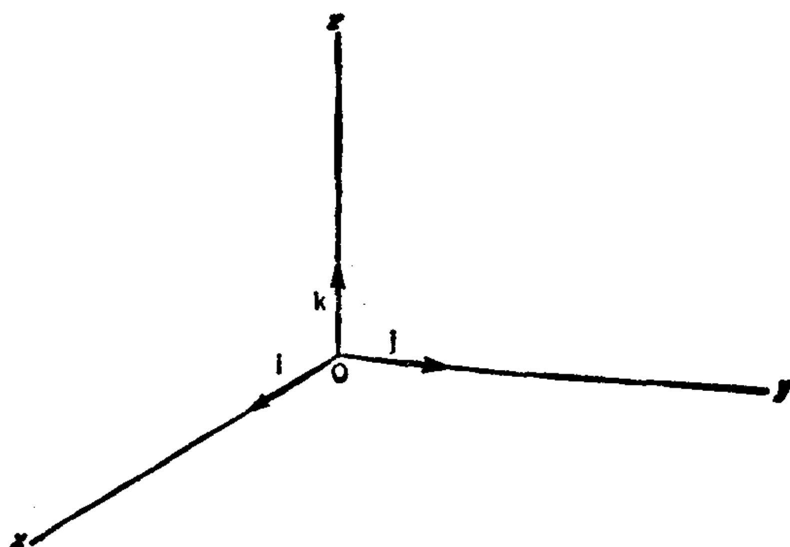


图 27 单位向量 **i**,**j** 和 **k**

$$\mathbf{k}\times\mathbf{i}=\mathbf{j}=-\mathbf{i}\times\mathbf{k},$$

$$\mathbf{i}\times\mathbf{i}=\mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=\mathbf{0}.$$

我们有

$$\mathbf{i}=(1,0,0),\quad \mathbf{j}=(0,1,0),\quad \mathbf{k}=(0,0,1).$$

验证是例行公事．例如

$$\mathbf{i}\times\mathbf{j}=(1,0,0)\times(0,1,0)=(0,0,1)=\mathbf{k}$$

和

$$\mathbf{j}\times\mathbf{i}=(0,1,0)\times(1,0,0)=(0,0,-1)=-\mathbf{k}.$$

$\mathbf{v}\times\mathbf{v}'$ 的长度

由公式(18)我们有

$$\begin{aligned}|\mathbf{v}\times\mathbf{v}'|^2&=(bc'-cb')^2+(ca'-ac')^2+(ab'-ba')^2\\&=(a^2+b^2+c^2)(a'^2+b'^2+c'^2)-(aa'+bb'+cc')^2\\&=|\mathbf{v}|^2|\mathbf{v}'|^2-(\mathbf{v}\cdot\mathbf{v}')^2\\&=|\mathbf{v}|^2|\mathbf{v}'|^2-(|\mathbf{v}||\mathbf{v}'|\cos\theta)^2\quad(\text{用 4.2 节公式(7)})\end{aligned}$$

因为

$$1-\cos^2\theta=\sin^2\theta,$$

可得

$$\|\mathbf{v}\times\mathbf{v}'\|=\|\mathbf{v}\|\|\mathbf{v}'\|\sin\theta, \tag{19}$$

这里θ是$\mathbf{v}$和$\mathbf{v}'$的夹角.

$\mathbf{v}\times\mathbf{v}'$的长度等于以$\mathbf{v}$和$\mathbf{v}'$为相邻边的平行四边形面积(图28).

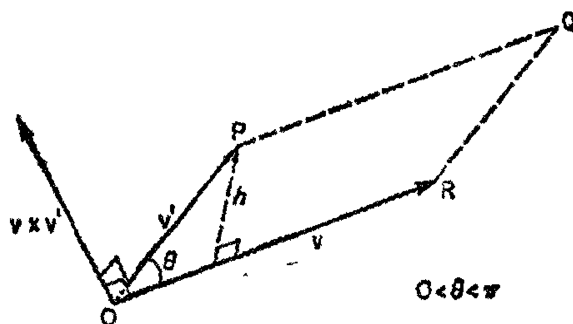


图 28 向量积

证明 面积$OPQR=OR\cdot h=\|\mathbf{v}\|\|\mathbf{v}'\|\sin\theta=\|\mathbf{v}\times\mathbf{v}'\|$.

注意对非零向量$\mathbf{v}$和$\mathbf{v}'$, $\mathbf{v}\times\mathbf{v}'=\mathbf{0}$当且仅当$\mathbf{v}$平行于$\mathbf{v}'$. 特别, $\mathbf{v}\times\mathbf{v}=\mathbf{0}$对一切$\mathbf{v}$成立.

$\mathbf{v}\times\mathbf{v}'$的方向

设在一个右手直角坐标中,$\mathbf{v}=(a,b,c)$和$\mathbf{v}'=(a',b',c')$是二个非零的互不平行的向量. 我们将证明向量$\mathbf{v}\times\mathbf{v}'$垂直于$\mathbf{v}$和$\mathbf{v}'$, 且$\mathbf{v}$(姆指),$\mathbf{v}'$(食指)和$\mathbf{v}\times\mathbf{v}'$(中指)构成一个右手三向量组(见图28).

证明 我们知道$\mathbf{v}\times\mathbf{v}'$是长度由公式(19)给出且垂直于$\mathbf{v}$和$\mathbf{v}'$的非零向量. 剩下要证明$\mathbf{v}$, $\mathbf{v}'$和$\mathbf{v}\times\mathbf{v}'$构成一个右手的而不是左手的三向量组.

我们用连续性来证明. 由公式(18), 显然连续改变$\mathbf{v}$和$\mathbf{v}'$导致连续改变$\mathbf{v}\times\mathbf{v}'$. 我们把$\mathbf{v}$连续地变到$\mathbf{i}$并把$\mathbf{v}'$连续地变到$\mathbf{j}$(图29), 使在变化过程的任何一步所得向量既不平行又不为零.

在这样的过程中, $\mathbf{v}\times\mathbf{v}'$总不取值$\mathbf{0}$并连续地变为$\mathbf{i}\times\mathbf{j}=\mathbf{k}$.

这样 $\mathbf{v},\mathbf{v}'$ 和 $\mathbf{v}\times\mathbf{v}'$ 也构成一个右手三向量组，因为把左手三向量组变为右手三向量组的连续过程要求叉积在某一步等于零.

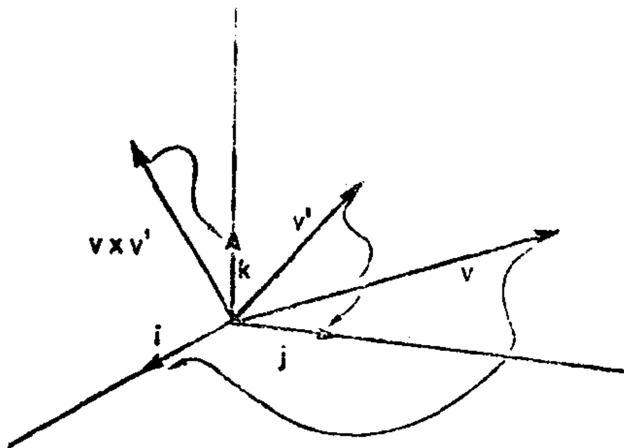


图 29 三向量组的连续变换

总结

设 $\mathbf{v}=(a,b,c)$ 和 $\mathbf{v}'=(a',b',c')$ 是由一个右手直角坐标系表示的向量. 定义

$$\mathbf{v}\times\mathbf{v}'=(bc'-cb',\ ca'-ac',ab'-ba').$$

于是：

命题 I

$$\|\mathbf{v}\times\mathbf{v}'\|=\|\mathbf{v}\|\|\mathbf{v}'\|\sin\theta,$$

这里 θ 是 $\mathbf{v}$ 和 $\mathbf{v}'$ 的夹角.

命题 II 如果 $\mathbf{v}$ 和 $\mathbf{v}'$ 不是零向量又不平行，则 $\mathbf{v}\times\mathbf{v}'$ 垂直于 $\mathbf{v}$ 和 $\mathbf{v}'$,而且使 $\mathbf{v},\mathbf{v}'$ 和 $\mathbf{v}\times\mathbf{v}'$ 构成一个右手三向量组.（在其余情形，由命题 I $\mathbf{v}\times\mathbf{v}'=\mathbf{0}$.）

常常把命题 I 和命题 II 作为向量积的定义，这个定义不用参照一个特殊的坐标系.

例 12 证明

$$\mathbf{v}\times\mathbf{v}'=-\mathbf{v}'\times\mathbf{v} \tag{20}$$

从定义(18)便直接得出这点.

向量积是一种非交换运算. 以后我们将碰到另一种重要的非交换运算——矩阵乘法.

例 13 验证向量积是双线性的：对一切向量 $\mathbf{v},\mathbf{w},\mathbf{s}$ 和纯量 k,l 有

$$\mathbf{s}\times(k\mathbf{v}+l\mathbf{w})=k(\mathbf{s}\times\mathbf{v})+l(\mathbf{s}\times\mathbf{w})$$

$$(k\mathbf{v}+l\mathbf{w})\times\mathbf{s}=k(\mathbf{v}\times\mathbf{s})+l(\mathbf{w}\times\mathbf{s}). \qquad (21)$$

利用定义(18)便可验证.

例 14 向量积是结合的吗?

不是. 我们用下面的例子来说明这一点:

$$(\mathbf{i}\times\mathbf{i})\times\mathbf{j}\neq\mathbf{i}\times(\mathbf{i}\times\mathbf{j}),$$

这是因为左边等于 $\mathbf{0}\times\mathbf{j}=\mathbf{0}$, 而右边等于 $\mathbf{i}\times\mathbf{k}=-\mathbf{j}$.

表达式 $\mathbf{a}\times\mathbf{b}\times\mathbf{c}$ 是有歧义的. 必须加括号指明以怎样的次序构成这个叉积.

以下修改了的结合律,即雅可比恒等式,对一切向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 成立:

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})+\mathbf{b}\times(\mathbf{c}\times\mathbf{a})+\mathbf{c}\times(\mathbf{a}\times\mathbf{b})=\mathbf{0}. \qquad (22)$$

可以用关于向量三重积的公式

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c} \qquad (23)$$

来验证雅可比恒等式,而要证明公式(23),可设

$$\mathbf{a}=(a_x,a_y,a_z),\quad \mathbf{b}=(b_x,b_y,b_z),\quad \mathbf{c}=(c_x,c_y,c_z),$$

再用向量积和点积的定义即可.

力学上的应用

力矩

以 $M=p\|\mathbf{F}\|$ 定义力 $\mathbf{F}$ 关于点 O 的力矩 M, 这里 p 是 O 到 $\mathbf{F}$ 作用线 L 的垂直距离(图 30). 设 Q 是 L 上任意一点, 令 $\overrightarrow{OQ}=\mathbf{r}$. 则

$$M=p\|\mathbf{F}\|=\|\mathbf{r}\|\|\mathbf{F}\|\sin\theta=\|\mathbf{r}\times\mathbf{F}\|.$$

向量 $\mathbf{M}=\mathbf{r}\times\mathbf{F}$ 称为 $\mathbf{F}$ 关于 O 的向量矩.

对于一个右手直角坐标系, 设 $\mathbf{F}=(F_x,F_y,F_z)$, $\mathbf{r}=(x,y,z)$. 于是由

$$M_x=yF_z-zF_y,\quad M_y=zF_x-xF_z,\quad M_z=xF_y-yF_x$$

给出 $\mathbf{M}$ 的分量.

角动量

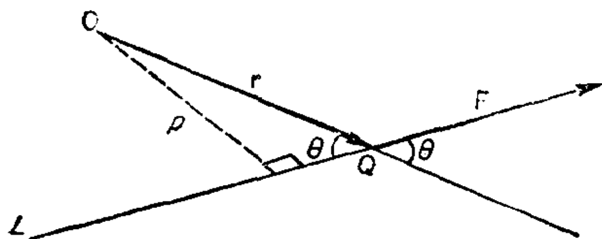


图 30 力矩

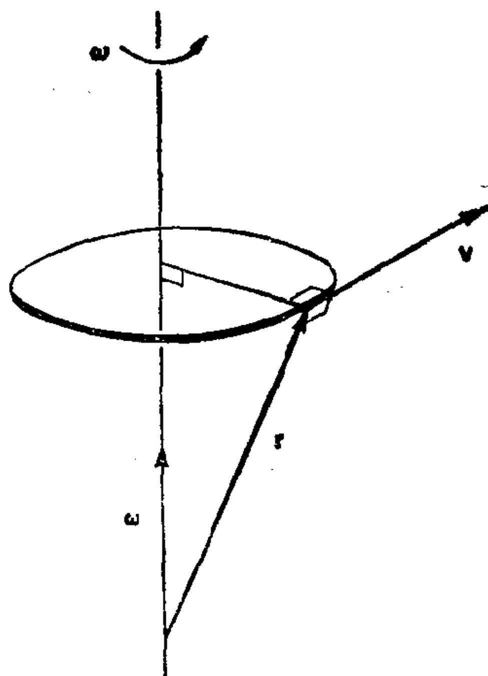


图 31 角速度

设想一个质量m的质点在图 30Q处. 假设这质点沿$\mathbf{F}$方向以速度$\mathbf{v}$运动. 于是把这质点关于O的动量矩(或称角动量)定义为向量$\mathbf{r}\times m\mathbf{v}$. 角动量的分量是

$$m(yv_z-zv_y),\quad m(zv_x-xv_z),\quad m(xv_y-yv_x).$$

角速度

例 15 一个刚体关于过点O的固定轴以角速度ω(弧度/秒)旋转. 确定这刚体上一点P的速度.

设$\boldsymbol{\omega}$是一个向量,其长度是ω,并指向右螺旋在所给旋转作用下的前进方向(图 31). 设$\mathbf{r}$是P关于O的位置向量. P沿着向量$\boldsymbol{\omega}\times\mathbf{r}$方向运动,向量$\boldsymbol{\omega}\times\mathbf{r}$垂直于包含$\boldsymbol{\omega}$和$\mathbf{r}$的平面. 进一步, P有速率

$$\omega p=\|\boldsymbol{\omega}\|\|\mathbf{r}\|\sin\theta=\|\boldsymbol{\omega}\times\mathbf{r}\|.$$

因此

$$\mathbf{v}=\boldsymbol{\omega}\times\mathbf{r}$$

给出P的速度. 向量$\boldsymbol{\omega}$称为刚体的角速度.

练　　习

25．求正交于$\mathbf{v}=(2,-1,3)$和$\mathbf{v}'=(1,3,-2)$的一个单位向量.

26．求包含$P(1,1,1)$,$Q(3,0,4)$和$R(2,4,-1)$的平面方程.

[提示：观察$\overrightarrow{PQ}=(2,-1,3)$和$\overrightarrow{PR}=(1,3,-2)$. 用练习 25 求出平面的单位法向量.]

27．求以练习 26 中P,Q,R为顶点的三角形面积.

[提示：考察以$\overrightarrow{PQ}$和$\overrightarrow{PR}$为边的平行四边形.]

28．如果$\mathbf{a}\times\mathbf{c}=\mathbf{b}\times\mathbf{c}$，可推出$\mathbf{a}=\mathbf{b}$吗？

29．证明：如果$\mathbf{a}\times\mathbf{c}=\mathbf{b}\times\mathbf{c}$对一切$\mathbf{c}$成立，则$\mathbf{a}=\mathbf{b}$.

30．用几何方法证明$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$是向量$\mathbf{b}$和$\mathbf{c}$所在平面中的一个向量.

31．当$\mathbf{a}=(1,-1,1)$,$\mathbf{b}=(1,2,0)$,$\mathbf{c}=(3,-2,1)$时，验证公式(23).

32．验证雅可比恒等式.

33．对$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$求出一个类似于(23)的公式.

[提示：应用公式(20).]

34．设$\mathbf{a},\mathbf{b},\mathbf{c}$是点$A,B,C$(关于原点$O$)的位置向量. 假设$\mathbf{a},\mathbf{b},\mathbf{c}$构成一个右手三向量组. 证明纯量三重积$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$等于以$OA,OB,OC$为边的平行六面体的体积. 由此推出

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})=\mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})=\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}).$$

然而，必须注意，

$$\mathbf{a}\cdot(\mathbf{c}\times\mathbf{b})=-\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}).$$

条件

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})=0$$

意味着什么？

35．一个力$\mathbf{F}=(1,-1,2)$的作用线通过点$Q(2,1,0)$. 求$\mathbf{F}$关于下列各点的力矩：

(i)　关于原点O；

(ii)　关于点$A(1,1,1)$；

(iii)　关于点$B(5,-2,6)$.

36．一个半径 3 cm 的球以角速率$\omega=6$弧度/秒绕过其中心O的固定轴旋转. 旋转轴在向量$\mathbf{v}=(1,2,2)$方向上，$\mathbf{v}$是由以O为原点的一个右手直角坐标系表示的. 如果你沿$\mathbf{v}$的方向看，旋转是顺时针方向的. 求球表面点

$A(3,0,0)$，$B(0,3,0)$，$C(0,0,3)$，$D(1,2,2)$，$E(2,1,-2)$的线速度．作一个草图．

37．一个刚体绕过O的轴旋转，旋转轴和一个直角坐标系的坐标轴Ox，Oy，Oz都构成相等的锐角，和通常一样，把Oz轴取作铅垂方向．在特定的一瞬时刚体上一点P有坐标$(3,3,4)$．证明在这一瞬时这一点的速度是水平方向的．

第五章 内 积 空 间

在第四章里，我们引进了二维或三维空间里的点积．我们已知道长度，角和正交性都可以用点积表达出来．在本章里，我们将推广这些概念．首先我们把点积概念适当推广．我们在 5.1 节中做这一点，在那里我们引进一种通常称为内积的纯量积，其代数性质类似于点积．点积是内积的一种特殊情形．

带内积的实向量空间称为欧几里得空间．在欧几里得空间中有正交以及长度的概念，并且欧几里得几何的许多定理（例如毕达哥拉斯定理）都有相应的解释（见 5.2 节例 7）．

二维和三维欧几里得几何中的定理总能用图形或模型来说明．然而我们可能已注意，当把这些定理推广到高维空间后，再要从几何上加以说明是很困难的了．推广了的正交与长度的概念的实质与其说是几何的，不如说是代数的．

结束本章时，我们将简短讨论复内积空间，即酉空间．

5.1 内积：欧几里得向量空间

设 $\mathscr{V}$ 是实数上的向量空间．如果对任何一对向量 $\mathbf{v}, \mathbf{w}$ 确定一个实数 $\langle \mathbf{v}, \mathbf{w} \rangle$ 满足*

(i) $\langle \mathbf{v},\mathbf{w} \rangle = \langle \mathbf{w},\mathbf{v} \rangle$（对称性）；

(ii) 对一切向量 $\mathbf{v},\mathbf{w},\mathbf{s}$ 和纯量 k,l，

$$\langle \mathbf{s}, k\mathbf{v}+l\mathbf{w} \rangle = k\langle \mathbf{s},\mathbf{v} \rangle + l\langle \mathbf{s},\mathbf{w} \rangle$$

$$\langle k\mathbf{v}+l\mathbf{w}, \mathbf{s} \rangle = k\langle \mathbf{v},\mathbf{s} \rangle + l\langle \mathbf{w},\mathbf{s} \rangle \text{（双线性）；}$$

(iii) 如果 $\mathbf{v} \neq \mathbf{0}$，则 $\langle \mathbf{v},\mathbf{v} \rangle > 0$，且 $\langle \mathbf{0},\mathbf{0} \rangle = 0$（正定性）；

---

* 和 4.2 节定理 I 比较．

则函数$\langle , \rangle$称为$\mathscr{V}$上的内积(或纯量积)，而$\mathscr{V}$称为欧几里得向量空间.

于是，要使实数上的一个向量空间成为欧几里得向量空间，必须在它上面定义一个实值内积. 由 4.2 节定理I，我们立刻可见带有普通点积的三维空间是一个欧几里得向量空间. 实际上，上述定义正是受这个定理的启发而作出的.

例 1　两个重要的欧几里得向量空间

(i) $\mathscr{R}^n$上定义内积为

$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n, \tag{1}$$

这里 $\mathbf{v} = (v_1, v_2, \cdots, v_n)$，$\mathbf{w} = (w_1, w_2, \cdots, w_n)$. 显然上述定义满足对称性，双线性和正定性这些条件. 我们将用$\mathscr{E}^n$表示这个欧几里得向量空间. 对 $n=2$ 或 3，$\mathscr{E}^n$ 的内积就是点积.

(ii) 在区间 $a \leqslant x \leqslant b$ 上连续的一切实值函数的向量空间上定义内积为

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx. \tag{2}$$

我们将用$\mathscr{F}_{a,b}$表示这个欧几里得向量空间.

让我们来验证定义(2)适合内积所要求的三个性质.

对称性：

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx = \int_a^b g(x) f(x) dx = \langle g, f \rangle.$$

双线性：

$$\begin{aligned}\langle kf + lg, h \rangle &= \int_a^b (kf + lg)(x) h(x) dx \\ &= \int_a^b (kf(x) + lg(x)) h(x) dx \\ &= k\int_a^b f(x) h(x) dx + l\int_a^b g(x) h(x) dx \\ &= k\langle f, h \rangle + l\langle g, h \rangle.\end{aligned}$$

由对称性得出另一个线性条件.

正定性：假设 f 是在区间 $a \leqslant x \leqslant b$ 上不恒等于零的一个连续

函数，则根据关于连续函数的一个著名定理

$$\langle f,f\rangle=\int_a^b(f(x))^2dx>0.$$

最后，如果 θ 是区间上的一个零函数，即对一切 x ($a\leqslant x\leqslant b$)，$\theta(x)=0$，则

$$\langle\theta,\theta\rangle=\int_a^b 0dx=0.$$

由(1)和(2)定义的内积有时称为它们相应向量空间上的标准内积.

例 2 已知一个欧几里得向量空间 $\mathscr{E}$ 具有零向量 θ，证明对 $\mathscr{E}$ 中一切 f 有

$$\langle\theta,f\rangle=0=\langle f,\theta\rangle.$$

因为 $f+\theta=f$，所以我们有

$$\begin{aligned}\langle f,f\rangle&=\langle f+\theta,f\rangle\\&=\langle f,f\rangle+\langle\theta,f\rangle.\end{aligned}$$

因此

$$\langle\theta,f\rangle=0.$$

用对称性，还有

$$\langle f,\theta\rangle=0.$$

练　　习

1．在欧几里得向量空间的定义中，证明第二个线性关系是第一个线性关系和对称性的推论.

2．以下哪些函数 $\langle,\rangle$ 定义了 $\mathscr{R}^3$ 上的一个内积？

(i) $\langle\mathbf{v},\mathbf{w}\rangle=v_1w_1+v_2w_2+v_3w_3+2v_2w_3-2v_3w_2$;

(ii) $\langle\mathbf{v},\mathbf{w}\rangle=v_1w_1+v_2w_2+v_3w_3-2v_2w_3-2v_3w_2$;

(iii) $\langle\mathbf{v},\mathbf{w}\rangle=v_1w_1+v_2w_2+v_3w_3-v_2w_3-v_3w_2$;

(iv) $\langle\mathbf{v},\mathbf{w}\rangle=v_1w_1+v_2w_2+v_3w_3-\frac{1}{2}v_2w_3-\frac{1}{2}v_3w_2$;

(v) $\langle\mathbf{v},\mathbf{w}\rangle=v_1^2w_1^2+v_2^2w_2^2+v_3^2w_3^2$.

3．以下哪些函数$\langle\ ,\ \rangle$定义了一切x的实多项式空间$\mathscr{P}$上的一个内积？

（i） $\langle f,g\rangle=\int_{-1}^{1}f(x)g(x)dx$；

（ii） $\langle f,g\rangle=\int_{-1}^{1}xf(x)g(x)dx$；

（iii） $\langle f,g\rangle=\int_{-1}^{1}x^2f(x)g(x)dx$；

（iv） $\langle f,g\rangle=\int_{0}^{1}xf(x)g(x)dx$；

（v） $\langle f,g\rangle=-\int_{0}^{1}xf(x)g(x)dx$；

（vi） $\langle f,g\rangle=\int_{0}^{\infty}e^{-x}f(x)g(x)dx$。

5.2 欧几里得向量空间中的正交性和长度

设$\mathscr{E}$是具有内积$\langle,\rangle$的欧几里得向量空间．设f和g是$\mathscr{E}$中的向量*．我们定义正交性和长度如下：

(i) f和g正交当且仅当

$$\langle f,g\rangle=0;$$

(ii) f的长度是：

$$\|f\|=\sqrt{\langle f,f\rangle},$$

$\|f\|$又称为f的模．

因为$\mathscr{E}$的内积是正定的，所以$\mathscr{E}$中每一个非零向量被赋予了一个正实数长度，并且零向量的长度是零，此外，对任何纯量k，

$$\|kf\|=|k|\|f\|.$$

以上定义推广了在$\mathscr{E}^2$和$\mathscr{E}^3$中通常正交性和长度的概念(见4.2节)．

注意正交性的定义关于f和g是对称的：f和g正交当且仅当g和f正交．我们也说f正交于g，或g正交于f．

---

* 因为许多欧几里得向量空间把函数作为它们的向量，所以我们放弃了在讨论一般理论时采用的粗体向量记号．

例 3 在$\mathscr{E}^5$中，求$\mathbf{v}=(2,-1,0,4,-2)$的长度，再证明$\mathbf{v}$正交于$\mathbf{w}=(3,4,2,1,3)$.

$$\|\mathbf{v}\|^2=\langle\mathbf{v},\mathbf{v}\rangle=4+1+0+16+4=25.$$

因此

$$\|\mathbf{v}\|=5.$$

而且

$$\langle\mathbf{v},\mathbf{w}\rangle=6-4+0+4-6=0.$$

因此$\mathbf{v}$正交于$\mathbf{w}$.

例 4 勒让德多项式

在$\mathscr{F}_{-1,1}$中证明，由勒让德多项式

$$P_1(x)=x,\quad P_2(x)=\frac{1}{2}(3x^2-1)$$

定义的函数P_1和P_2是正交的．求它们的长度.

$$\langle P_1;P_2\rangle=\int_{-1}^{1}P_1(x)P_2(x)dx=\frac{1}{2}\int_{-1}^{1}(3x^3-x)dx=0$$

因此P_1和P_2是正交的.

$$\|P_1\|^2=\langle P_1,P_1\rangle=\int_{-1}^{1}x^2dx=\left[\frac{1}{3}x^3\right]_{-1}^{1}=\frac{2}{3}.$$

因此

$$\|P_1\|=\sqrt{\frac{2}{3}}.$$

类似地可证明

$$\|P_2\|=\sqrt{\frac{2}{5}}.$$

勒让德多项式在物理学中比较重要．例如它们出现在拉普拉斯方程和薛定谔方程（1.3 节例 4）的用球极坐标表出的解中．这些多项式的一般定义和进一步的探讨可见 5.5 节例 18.

例 5 三角函数

在 $\mathscr{F}_{-\pi,\pi}$ 中证明下列三角函数相互正交并且它们的长度都等于1：

$$S_n(n=1,2,3,\cdots); \quad C_m(m=0,1,2,3,\cdots);$$

这里

$$S_n(x)=\frac{1}{\sqrt{\pi}}\sin nx,$$

$$C_m(x)=\frac{1}{\sqrt{\pi}}\cos mx \qquad (m\neq 0),$$

$$C_0(x)=\frac{1}{\sqrt{2\pi}}.$$

我们先证明当 $n\neq k$ 时 S_n 和 S_k 正交及 $\|S_n\|=1$。

$$\langle S_n, S_k\rangle=\frac{1}{\pi}\int_{-\pi}^{\pi}\sin nx\sin kx\,dx$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}(\cos(n-k)x+\cos(n+k)x)dx,$$

因此，当 $n\neq k$ 时，

$$\langle S_n, S_k\rangle=\frac{1}{2\pi}\left[\frac{\sin(n-k)x}{n-k}+\frac{\sin(n+k)x}{n+k}\right]_{-\pi}^{\pi}=0,$$

而当 $n=k$ 时，

$$\|S_n\|^2=\langle S_n,S_n\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}1\,dx=1,$$

因此

$$\|S_n\|=1.$$

用类似的方法可证

$$\|C_m\|=1 \qquad (m=0,1,2,\cdots).$$

(注意

$$\|C_0\|^2=\int_{-\pi}^{\pi}\frac{1}{2\pi}dx=1\ )$$

以及对 $m\neq k$

$$\langle C_m, C_k\rangle=0.$$

还有对一切m和n

$$\langle C_m, S_n\rangle=0.$$

在区间

$$-\pi\leqslant x\leqslant\pi$$

上余弦和正弦函数的正交性，是傅里叶分析理论的基础。

奇函数和偶函数

计算内积常常要涉及形如

$$\int_{-l}^{l} f(x)dx$$

的积分，积分区间是关于原点对称的．有一类重要的函数，它们的积分取零值．函数 f 称为奇函数，如果对一切 x

$$f(-x)=-f(x).$$

函数 $\sin nx$ 是奇函数；对奇数 r，x^r 也是奇函数．我们在图 32 (a)中画出了另一个例子．

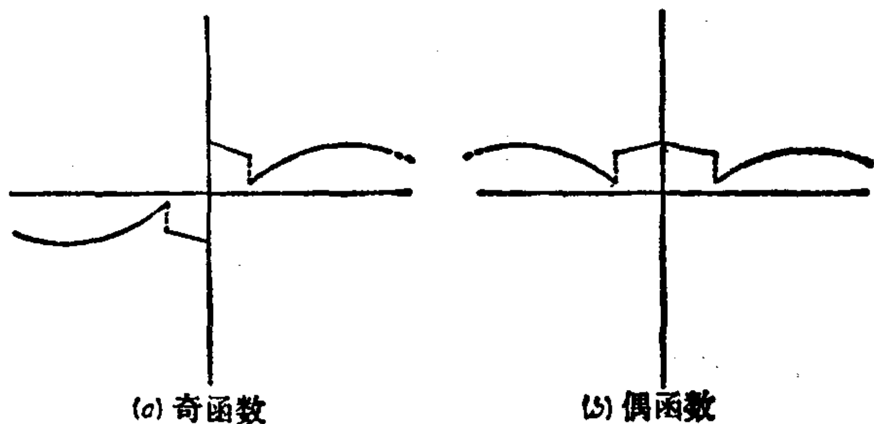


图 32 奇函数和偶函数

显然，如果 f 是一个奇函数，则

$$\int_{-l}^{l} f(x)dx=0,$$

因为$\int_{-l}^{0} f(x)dx$ 的值是$\int_{0}^{l} f(x)dx$ 的相反数.

以上结果可以用来求例 5 中的$\langle C_m, S_n\rangle$ 的值. 不必再计算，只要观察到

$$f(x)=\cos mx \sin nx$$

确定了一个奇函数，便得$\langle C_m, S_n\rangle=0$.

另外一类重要的函数是偶函数. 如果对一切 x 有

$$f(-x)=f(x),$$

则 f 是偶函数. 函数 $\cos nx$ 是偶函数，对偶数 r, x^r 也是偶函数. 图 32 (b)是另一个偶函数的例子. 如果 f 是偶函数，则

$$\int_{-l}^{l} f(x)dx=2\int_{0}^{l} f(x)dx.$$

计算这种积分一般没有什么捷径.

大多数函数既不是偶函数也不是奇函数

(例如：$1+x$, $\sin x+\cos x$, e^x).

例 6 埃尔米特多项式

证明埃尔米特多项式

$$H_0(x)=1,\quad H_1(x)=2x,\quad H_2(x)=4x^2-2$$

确定的函数 H_0, H_1, H_2 关于内积

$$\langle f, g\rangle=\int_{-\infty}^{\infty} e^{-x^2} f(x) g(x) dx \tag{3}$$

是互相正交的.

我们先考察

$$\langle H_0, H_1\rangle=\int_{-\infty}^{\infty} 2x e^{-x^2} dx.$$

被积函数是一个奇函数. 因此$\langle H_0, H_1\rangle=0$. 类似地，$\langle H_1, H_2\rangle=0$. 最后

$$\langle H_0, H_2\rangle=\int_{-\infty}^{\infty}(4x^2-2) e^{-x^2} dx.$$

已知

$$\int_{-\infty}^{\infty} x^{2n}e^{-x^2}dx=\begin{cases}\sqrt{\pi} & (当\ n=0)\\ \dfrac{1\cdot3\cdot5\cdots(2n-1)}{2^n}\cdot\sqrt{\pi} & (当\ n=1,2,3,\cdots),\end{cases} \tag{4}$$

由此得出$\langle H_0,H_2\rangle=0$.

在某些物理学问题中，如用量子力学来讨论简谐振子时，会出现埃尔米特多项式．此类多项式的一般定义与进一步的探讨可参看5.5节例19.

因为由(3)定义的内积的积分区域是无限的，所以某些连续函数关于这个内积有无限的长度(例如，取$f(x)=e^{x^2}$)．对一切x连续的所有函数的集合，对内积(3)不构成一个欧几里得向量空间．在欧几里得空间里，两个向量的内积必须是一个实数，这里是排除无穷大的.

可以证明，所有关于内积(3)具有有限长度的函数集合构成一个欧几里得向量空间(见5.3节练习19).

我们把欧几里得几何中一个熟知的定理推广来结束本节.

例 7 毕达哥拉斯定理及其逆

证明 欧几里得空间中两个向量f和g正交当且仅当

$$\|f+g\|^2=\|f\|^2+\|g\|^2.$$

$$\begin{aligned}\|f+g\|^2&=\langle f+g,f+g\rangle\\&=\langle f,f\rangle+\langle f,g\rangle+\langle g,f\rangle+\langle g,g\rangle\\&\quad(\text{利用内积的双线性}),\\&=\|f\|^2+\|g\|^2+2\langle f,g\rangle.\end{aligned}$$

因此

$$\|f+g\|^2=\|f\|^2+\|g\|^2$$

当且仅当$\langle f,g\rangle=0$，即当且仅当f正交于g.

这便推广了毕达哥拉斯定理(见4.2节练习12).

练　　习

4．在$\mathscr{E}^4$中，验证向量$(1,1,1,1)$，$(1,-1,1,-1)$，$(1,1,-1,-1)$和$(1,-1,-1,1)$是互相正交的．求它们的长度．

5．证明，$\mathscr{F}_{-1,1}$中任意两个形如

$$a_0+a_2x^2+\cdots+a_{2n}x^{2n} \text{ 和 } a_1x+a_3x^3+\cdots+a_{2n+1}x^{2n+1}$$

的多项式互相正交．

6．在$\mathscr{F}_{-1,1}$中验证勒让德多项式

$$P_0(x)=1 \text{ 和 } P_3(x)=\frac{1}{2}(5x^3-3x)$$

是相互正交的．

7．验证埃尔米特多项式H_0,H_1和H_2在$\mathscr{F}_{-1,1}$中(即对于内积

$$\langle f, g\rangle=\int_{-1}^{1} f(x)g(x)dx)$$

不是互相正交的．

8．验证埃尔米特多项式

$$H_3(x)=8x^3-12x$$

对于例6的内积(3)，正交于H_0,H_1和H_2．

9．证明拉盖尔多项式

$$L_0(x)=1,\quad L_1(x)=1-x,\quad L_2(x)=2-4x+x^2$$

对于内积

$$\langle f, g\rangle=\int_{0}^{\infty} e^{-x}f(x)g(x)dx$$

是相互正交的．求关于这个内积的上述多项式的长度．

10．验证，对一切m和$n(n\neq 0)$，$\cos mx \sin nx$是一个奇函数．

11．证明，两个偶函数的积是偶函数，两个奇函数的积是偶函数，一个偶函数和一个奇函数的积是奇函数．

12．证明，如果f是偶函数，则

$$\int_{-\pi}^{\pi} f(x)\sin kx\, dx=0,$$

如果f是奇函数，则

$$\int_{-\pi}^{\pi} f(x)\cos kx\, dx=0.$$

13. 推广 4.2 节练习 11 和 13.

14. 完成例 5 的解.

5.3 一些重要的不等式

许瓦尔兹不等式

在 4.2 节里我们看到三维向量适合关系

$$\mathbf{v}\cdot\mathbf{w}=\|\mathbf{v}\|\|\mathbf{w}\|\cos\theta,$$

这里 θ 表示 v 和 w 的夹角.

因为 $|\cos\theta|\leqslant 1$ 对一切 θ 成立,所以我们立刻推得

$$|\mathbf{v}\cdot\mathbf{w}|\leqslant\|\mathbf{v}\|\|\mathbf{w}\|, \tag{5}$$

等式当且仅当 $\cos\theta=\pm 1$ 时成立,或者说当且仅当 v 和 w 线性相关时成立. 关系式(5)是一个著名的代数不等式,若把它两边分别平方并且用分量写出,则成为

$$(v_1w_1+v_2w_2+v_3w_3)^2\leqslant(v_1^2+v_2^2+v_3^2)(w_1^2+w_2^2+w_3^2),$$

它对任意实数 v_i, w_i 都成立. 现在我们来推广(5). 因为在一般的欧几里得空间没有定义两个向量夹角的余弦(我们仅仅定义了正交性), 所以为了推广这个不等式, 我们必须有一个代数的证法.

定理 I 许瓦尔兹不等式

在任意欧几里得向量空间中, 有

$$|\langle f,g\rangle|\leqslant\|f\|\|g\|, \tag{6}$$

等式成立的充要条件是 f 和 g 线性相关.

证明 我们先处理 f 和 g 线性相关的情形. 于是或有 $f=kg$ 对某纯量 k 成立,或有 $g=\theta$,即 g 是零向量. 如果是前者, 则

$$\|f\|=|k|\|g\|$$

且

$$|\langle f,g\rangle|=|\langle kg,g\rangle|=|k|\|g\|^2=\|f\|\|g\|.$$

如果 $g=\theta$,则显然不等式(6)两边都等于零.

以下假设 f 和 g 是线性无关的．此时 g 不是零向量，且对所有纯量 k，$f-kg$ 有正长度．考察 $f-kg$ 的长度：

$$\begin{aligned}0<\|f-kg\|^2&=\langle f-kg,f-kg\rangle,\\&=\langle f,f\rangle-k\langle g,f\rangle-k\langle f,g\rangle+k^2\langle g,g\rangle\\&=\|f\|^2-2k\langle f,g\rangle+k^2\|g\|^2,\end{aligned}$$

这里用了内积的对称性．

用 $\|g\|^2$ 乘这个不等式并且整理一下，我们得到

$$2k\|g\|^2\langle f,g\rangle-k^2\|g\|^4<\|f\|^2\|g\|^2.$$

这个不等式对一切纯量 k 成立．现在取 k 使

$$k\|g\|^2=\langle f,g\rangle.$$

这就得出

$$\langle f,g\rangle^2<\|f\|^2\|g\|^2.$$

取算术根，我们便得到带着严格不等号的不等式（6），证明便完成了．

许瓦尔兹不等式有许多应用．这里举几个例子．

例 8　柯西不等式

证明柯西不等式：对任意实数 $x_1,\cdots,x_n$ 和 $y_1,\cdots,y_n$，

$$(x_1y_1+\cdots+x_ny_n)^2\leqslant(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)\qquad(7)$$

等号当且仅当向量 $(x_1,\cdots,x_n)$ 和 $(y_1,\cdots,y_n)$ 线性相关时成立．

把许瓦尔兹不等式直接用到 $\mathscr{E}^n$ 上，就是柯西不等式．

柯西不等式在研究无穷级数时是相当重要的．例如，假设我们已知 $\sum_{i=1}^{\infty}x_i^2$ 和 $\sum_{i=1}^{\infty}y_i^2$ 都是收敛的．从不等式(7)得出 $\sum_{i=1}^{\infty}|x_iy_i|$ 也是收敛的．

例 9　积分的许瓦尔兹不等式

f 和 g 是区间 $a\leqslant x\leqslant b$ 上的连续函数．证明

$$\left(\int_a^b f(x)g(x)dx\right)^2\leqslant\int_a^b(f(x))^2dx\int_a^b(g(x))^2dx,\qquad(8)$$

等式当且仅当 f 和 g 线性相关时成立.

把许瓦尔兹不等式用于 $\mathscr{F}_{a,b}$,即得结论.

当某个积分无法准确求值时，若要找出该积分上界的一个近似值,不等式(8)是很有用的.

例 10 证明

$$\int_0^1 \sqrt{(1+x^3)}\,dx < \sqrt{\frac{5}{4}}.$$

在不等式(8)中，令 $f(x)=\sqrt{(1+x^3)}$ 及 $g(x)=1$. 函数 f 和 g 是线性无关的. 因此我们得到严格不等式

$$\left(\int_0^1 \sqrt{(1+x^3)}\,dx\right)^2 < \int_0^1 (1+x^3)\,dx \int_0^1 1\,dx = \frac{5}{4}.$$

我们取算术平方根，解答便完成了.

我们还看到

$$\int_0^1 \sqrt{(1+x^3)}\,dx > \int_0^1 1\,dx = 1,$$

于是，对 $\int_0^1 \sqrt{(1+x^3)}\,dx$ 有一个满意的估计：它的值在 1 和 1.1181 之间.

例 11 统计学上的应用：相关

统计学中许多问题需要同时考虑两个变量. 例如，你可能对一群人员的高度和重量感兴趣. 在这里变量分别是相应于高度和重量的 h 和 w. 自然会问 h 和 w 之间是否存在某种联系,以及怎样才能度量这种联系.

假设这群人有 n 个成员,分别用标号 $1,2,\cdots,n$ 表示，还假设第 i 个成员高 h_i,重 w_i. 我们可把测量得到的高度和重量分别用 $\mathscr{E}^n$ 中的向量 $\mathbf{h}=(h_1,\cdots,h_n)$ 与 $\mathbf{w}=(w_1,\cdots,w_n)$ 表示. 由定理 I,

$$|\langle\mathbf{h},\mathbf{w}\rangle| \leqslant \|\mathbf{h}\|\|\mathbf{w}\|,$$

等号当且仅当 **h** 和 **w** 间有一个线性关系时才成立。

在实际问题中，常常用与相应平均值的差异来度量这些高度和重量．h 的平均值定义为

$$\bar{h}=\sum_{i=1}^{n}\frac{1}{n}h_i.$$

类似地定义 w 的平均值．以上述高度和重量与相应平均值的差作为分量的两个向量

$$\mathbf{h}^*=(h_1-\bar{h},h_2-\bar{h},\cdots,h_n-\bar{h})$$

和

$$\mathbf{w}^*=(w_1-\bar{w},w_2-\bar{w},\cdots,w_n-\bar{w})$$

就可以用于记录这些高度和重量．我们有

$$\|\mathbf{h}^*\|^2=\sum_{i=1}^{n}(h_i-\bar{h})^2=n\sum_{i=1}^{n}\frac{1}{n}(h_i-\bar{h})^2=ns_h^2.$$

s_h^2 称为 h 的方差．类似地，

$$\|\mathbf{w}^*\|^2=ns_w^2.$$

$\mathbf{h}^*$ 和 $\mathbf{w}^*$ 的内积是

$$\langle\mathbf{h}^*,\mathbf{w}^*\rangle=n\sum_{i=1}^{n}\frac{1}{n}(h_i-\bar{h})(w_i-\bar{w})=ns_{hw}.$$

s_{hw} 称为 h 和 w 的协方差．由许瓦尔兹不等式

$$|s_{hw}|\leqslant s_h s_w,$$

即

$$-1\leqslant\frac{s_{hw}}{s_h s_w}\leqslant 1,$$

只要 s_h 和 s_w 都不等于零．（如果 $s_h=0$，则 $h_i=\bar{h}$ 对一切 i 成立，便有 $s_{hw}=0$．类似地，如果 $s_w=0$，则 $s_{hw}=0$。）

这个没有单位的数

$$r_{hw}=\frac{s_{hw}}{s_h s_w}$$

称为相关系数. 它适合不等式

$$-1\leqslant r_{hw}\leqslant 1,$$

而且当且仅当 h 和 w 线性相关时 r_{hw} 才取一端的值. 如果存在线性关系

$$w-\overline{w}=c(h-\overline{h}),$$

那末, c 是正常数或负常数取决于 r_{hw} 等于 1 或 -1. ($c=0$ 的情形对应于 $s_w=0=s_{hw}$.) 把这些点 (h_i,w_i) 画到一张图上, 所有点都在一条斜率为 c 的直线上.

当 $|r_{hw}|<1$, 这些点 (h_i,w_i) 不都在一条直线上. 然而, $|r_{hw}|$ 的值越大, 由这些点构成的分布图越接近于一条直线.

相关问题的进一步讨论可参阅 Paul G. Hoel, *Introduction to Mathematical Statistics*, Third Edition, Wiley, 1962.

三角形不等式

三维向量一个显然的不等式是三角形不等式

$$\|\mathbf{v}+\mathbf{w}\|\leqslant\|\mathbf{v}\|+\|\mathbf{w}\|.$$

它指出一个三角形两边长度之和不小于第三边的长度. 以下定理中我们证明这个不等式的一个推广.

定理 II 三角形不等式

对一个欧几里得空间 $\mathscr{E}$ 中一切 f,g 有

$$\|f+g\|\leqslant\|f\|+\|g\|. \tag{9}$$

证明
$$\begin{aligned}\|f+g\|^2&=\langle f+g,f+g\rangle\\&=\langle f,f\rangle+\langle g,f\rangle+\langle f,g\rangle+\langle g,g\rangle\\&=\|f\|^2+2\langle f,g\rangle+\|g\|^2\\&\leqslant\|f\|^2+2|\langle f,g\rangle|+\|g\|^2 \text{(根据许瓦尔兹不等式)}\\&=(\|f\|+\|g\|)^2.\end{aligned}$$

取算术平方根我们得到不等式

$$\|f+g\|\leqslant\|f\|+\|g\|,$$

于是完成了证明.

注意定理 II 的一个直接推论是不等式

$$\|kf+lg\|\leqslant|k|\|f\|+|l|\|g\|, \tag{10}$$

对一切纯量 k,l 成立.

如果我们把这个不等式用于 $\mathscr{F}_{a,b}$,便得到

$$\left(\int_a^b(kf(x)+lg(x))^2dx\right)^{\frac{1}{2}}\leqslant|k|\left(\int_a^b(f(x))^2dx\right)^{\frac{1}{2}}$$

$$+|l|\left(\int_a^b(g(x)^2dx\right)^{\frac{1}{2}}. \tag{11}$$

距离

我们推广距离概念(见 4.1 节例 2),定义一个欧几里得向量空间中两个向量 f 和 g 之间的距离为

$$\mathrm{d}(f,g)=\|f-g\|. \tag{12}$$

下列性质是距离的特征.

(i) $\mathrm{d}(f,g)=\mathrm{d}(g,f)$;

(ii) 当 $f\neq g$, $\mathrm{d}(f,g)>0$, $\mathrm{d}(f,f)=0$;

(iii) $\mathrm{d}(f,g)\leqslant\mathrm{d}(f,h)+\mathrm{d}(h,g)$. (13)

最后一个不等式是三角形不等式的变形,因为可把(13)写成如下形式

$$\|f-g\|\leqslant\|f-h\|+\|h-g\|, \tag{14}$$

而这不过是三角形不等式(9)用 $f-h$ 代替 f,用 $h-g$ 代替 g 而已.

练　　习

15. 用许瓦尔兹不等式证明

(i) $\int_0^\pi\sqrt{x\sin x}\,dx<\pi$;

(ii) $\int_0^\pi \sqrt{\sin x}\ dx < \sqrt{2\pi}$.

16. (霍伊尔)对以下12个大学生的身高(h)和体重(w)数据计算相关系数。

h(英尺)	63	72	70	68	66	69	74	70	63	72	65	71
w(磅)	124	184	161	164	140	154	210	164	126	172	133	150

画一个分布图。

17. 用一个草图来说明三角形不等式

$$\|\mathbf{v}+\mathbf{w}\| \leqslant \|\mathbf{v}\| + \|\mathbf{w}\|.$$

18. 在什么条件下三角形不等式是一个等式?

19. 利用不等式(10)证明,如果

$$\int_{-\infty}^{\infty} e^{-x^2}(f(x))^2 dx \quad 和 \quad \int_{-\infty}^{\infty} e^{-x^2}(g(x))^2 dx$$

都是有限的,则对任意纯量k, l

$$\int_{-\infty}^{\infty} e^{-x^2}(kf(x)+lg(x))^2 dx$$

也是有限的.(这个结果在研究埃尔米特多项式时是很重要的. 参看5.2节例6.)

20. 用一个草图来说明不等式

$$\|\mathbf{p}-\mathbf{q}\| \leqslant \|\mathbf{p}-\mathbf{r}\| + \|\mathbf{r}-\mathbf{q}\|,$$

这里$\mathbf{p}, \mathbf{q}$和$\mathbf{r}$依次表示点P, Q和R的位置向量.

21. 在一个欧几里得空间中,不等式

$$\|f-g\| \leqslant \|f\| - \|g\|$$

对所有向量f与g都成立吗?

5.4 正交集

直角坐标系与斜坐标系相比,其优越性在几何里是明显的. 相应地,要描述一个欧几里得向量空间,最好采用由相互正交的单位向量构成的基. 本节的目的是,证明在有限生成的欧几里得向

量空间里，存在这种基.

定义 I 相互正交的非零向量的一个集合称为一个正交集. 如果其中所有向量是单位向量，则这正交集称为一个正规正交集. 一个生成 $\mathscr{E}$ 的正交集称为 $\mathscr{E}$ 的一组正交基.

在上述定义中，"正规"一词指的是所涉及的向量都是单位(正规化的)向量.

例 12 向量 $\mathbf{e}_1=(1,0),\mathbf{e}_2=(0,1)$ 构成 $\mathscr{E}^2$ 的一组正规正交基.

$$\mathbf{e}_1^*=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),\qquad \mathbf{e}_2^*=\left(-\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)$$

也构成 $\mathscr{E}^2$ 的一组正规正交基.

在 $\mathscr{E}^3$ 中，$\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0\right),\left(-\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0\right)$ 构成一个正规正交集，然而却不构成一组基. $(1,1,0)$，$(-1,1,0)$，$(0,0,1)$ 在 $\mathscr{E}^3$ 中构成一个正交集，也构成一组基，但不构成一组正规正交基.

例 13 $f_1,\cdots,f_m$ 构成一个欧几里得向量空间中的一个正交集. 证明 $f_1,\cdots,f_m$ 线性无关.

假设向量满足一个线性关系

$$c_1f_1+\cdots+c_mf_m=\theta. \tag{15}$$

则对 $j=1,\cdots,m$，都有

$$\langle c_1f_1+\cdots+c_mf_m,f_j\rangle=\langle\theta,f_j\rangle=0$$

(用 5.1 节例 2).

因此

$$\begin{aligned}0&=c_1\langle f_1f_j\rangle+\cdots+c_m\langle f_m,f_j\rangle\\&=c_j\langle f_j,f_j\rangle,\quad 因为\langle f_i,f_j\rangle=0\quad(i\neq j),\\&=c_j\|f_j\|^2.\end{aligned}$$

而 $\|f_j\|\neq0$. 得出 $c_j=0$，因此关系(15)是一个平凡关系. 可见

$f_1,\cdots,f_m$ 是线性无关的.

例 14 构造正规正交基.

在 $\mathscr{E}^3$ 中求包含正交单位向量

$$\mathbf{u}_1=\left(\frac{1}{3},\ \frac{2}{3},\ -\frac{2}{3}\right) \quad 和 \quad \mathbf{u}_2=\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)$$

的正规正交基.

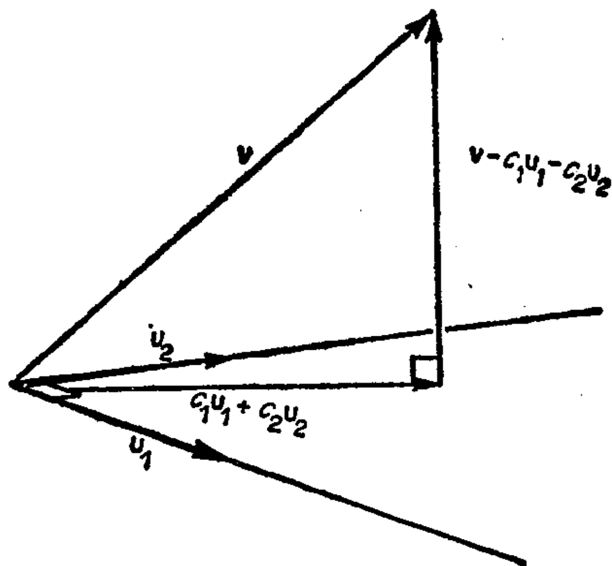


图 33 构造正规正交基

取任一个不是 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 线性组合的向量 $\mathbf{v}$. 这种取法是可能的，因为 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 生成 $\mathscr{E}^3$ 的一个真子空间(一个平面)(图 33).

取了v之后，再选纯量 c_1 和 c_2，使得 $\mathbf{v}-c_1\mathbf{u}_1-c_2\mathbf{u}_2$ 和 $\mathbf{u}_1$，$\mathbf{u}_2$ 都正交.这便要求

$$(\mathbf{v}-c_1\mathbf{u}_1-c_2\mathbf{u}_2)\cdot\mathbf{u}_1=0$$

和

$$(\mathbf{v}-c_1\mathbf{u}_1-c_2\mathbf{u}_2)\cdot\mathbf{u}_2=0.$$

记住

$$\mathbf{u}_2\cdot\mathbf{u}_1=\mathbf{u}_1\cdot\mathbf{u}_2=0 \quad 及 \quad \mathbf{u}_k\cdot\mathbf{u}_k=\|\mathbf{u}_k\|^2=1 \quad (k=1,2)$$

我们得到

$$c_k=\mathbf{v}\cdot\mathbf{u}_k \quad (k=1,2).$$

因此

$$\mathbf{w}=\mathbf{v}-\sum_{k=1}^{2}(\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k \tag{16}$$

和 $\mathbf{u}_1,\mathbf{u}_2$ 都正交. 我们还有 $\mathbf{w}\neq 0$，因为 $\mathbf{v}$ 不是 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 的线性组合. 最后，我们正规化 $\mathbf{w}$. 向量 $\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3=\frac{1}{\|\mathbf{w}\|}\mathbf{w}$ 构成 $\mathscr{E}^3$ 的一

组正规正交基.

我们已经抽象地描述了这个方法. 现在让我们用数字填上. 例如我们选 $\mathbf{v}=(1,0,0)$, 则

$$c_1=\mathbf{v}\cdot\mathbf{u}_1=\frac{1}{3} \quad 及 \quad c_2=\mathbf{v}\cdot\mathbf{u}_2=\frac{2}{3}.$$

因此

$$\mathbf{w}=(1,0,0)-\frac{1}{3}\left(\frac{1}{3},\ \frac{2}{3},\ -\frac{2}{3}\right)-\frac{2}{3}\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)$$

$$=\frac{1}{9}(4,-4,-2).$$

最后,我们令

$$\mathbf{u}_3=\frac{1}{\|\mathbf{w}\|}\mathbf{w}=\left(\frac{2}{3},\ -\frac{2}{3},\ -\frac{1}{3}\right).$$

例 14 有第二个解: 向量 $\mathbf{u}_1,\mathbf{u}_2$ 和 $-\mathbf{u}_3$ 也构成 $\mathscr{E}^3$ 的一组正规正交基. 除了这两组基, $\mathscr{E}^3$ 再没有别的包含 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 的正规正交基了.

现在我们来推广以上的例子.

例 15 格拉姆-施密特正交化过程

设 $f_1,\cdots,f_m$ 构成欧几里得空间 $\mathscr{E}$ 的一个正规正交集. 若 $f_1,\cdots,f_m$ 生成 $\mathscr{E}$ 的一个真子空间,证明存在一个单位向量 f_{m+1},使 $f_1,\cdots,f_m,f_{m+1}$ 构成一个正规正交集.

取不是 $f_1,\cdots,f_m$ 线性组合的任意一个向量 f. 这种取法是可能的,因为 $f_1,\cdots,f_m$ 生成 $\mathscr{E}$ 的一个真子空间. 现在选纯量 $c_1,\cdots,c_m$ 使得 $f-c_1f_1-\cdots-c_mf_m$ 正交于每个 $f_k(k=1,\cdots,m)$,也就是

$$\langle f-c_1f_1-\cdots-c_mf_m,f_k\rangle=0 \quad (k=1,\cdots,m),$$

由此

$$\langle f,f_k\rangle-c_1\langle f_1,f_k\rangle\cdots-c_m\langle f_m,f_k\rangle=0.$$

而

$$\langle f_i, f_k \rangle = 0 \quad (i \neq k), \quad 且 \quad \langle f_k, f_k \rangle = \| f_k \|^2 = 1.$$

因此

$$c_k = \langle f, f_k \rangle.$$

向量

$$g = f - \sum_{k=1}^{m} \langle f, f_k \rangle f_k \tag{17}$$

正交于每个 $f_k (k=1, \cdots, m)$. 而且, g 不是零向量, 因为 f 不是 $f_1, \cdots, f_m$ 的线性组合. 最后, 正规化 g, 得出所需向量

$$f_{m+1} = \frac{1}{\| g \|} g.$$

定理 III 任何有限维欧几里得向量空间 $\mathscr{E}$ 有一组正规正交基.

证明 选 $\mathscr{E}$ 的任意一个单位向量 f_1 作为基的第一个向量, 如果 $\mathscr{E}$ 是一维的, 则证明结束. 如果 $\mathscr{E}$ 是 n 维的 $(n>1)$, 利用格拉姆-施密特正交化过程 $n-1$ 次, 便得到一个正规正交集 $f_1, \cdots, f_n$. 这些向量是线性无关的(例 13), 因此构成 $\mathscr{E}$ 的一组基(3.3 节定理 II (a)).

我们可以把定理 III 加强一点.

定理 IV n 维欧几里得空间的任意 $m (<n)$ 个向量的正规正交集, 都可以扩充为一组正规正交基.

证明 应用格拉姆-施密特正交化过程 $n-m$ 次.

例 16 把正规正交集

$$\mathbf{u}_1 = \frac{1}{2}(1,1,1,1), \mathbf{u}_2 = \frac{1}{2}(1,-1,1,-1), \mathbf{u}_3 = \frac{1}{2}(1,1,-1,-1),$$

扩充为 $\mathscr{E}^4$ 的一组正规正交基.

如果你研究了 $\mathbf{u}_1, \mathbf{u}_2$ 和 $\mathbf{u}_3$ 中负号的分布, 也许你会准确地猜中一个单位向量 $\mathbf{u}_4$, 使它和上面三个向量一起成为 $\mathscr{E}^4$ 的正规正交基. 但是, 我们还是用这个例子来进一步说明格拉姆-施密特正

交化过程吧.

我们先取向量 $\mathbf{v}$，使 $\mathbf{v}$ 不是 $\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3$ 的线性组合. 我们可取 $\mathbf{v}=(1,0,0,0)$吗？我们暂时假定这样选择是合适的.

根据例 15 公式(17),向量

$$\mathbf{w}=\mathbf{v}-\sum_{k=1}^{3}\langle\mathbf{v},\mathbf{u}_k\rangle\mathbf{u}_k$$

正交于 $\mathbf{u}_1,\mathbf{u}_2$ 和 $\mathbf{u}_3$. 现在

$$\langle\mathbf{v},\mathbf{u}_1\rangle=\left\langle(1,0,0,0),\ \left(\frac{1}{2},\ \frac{1}{2},\ \frac{1}{2},\ \frac{1}{2}\right)\right\rangle=\frac{1}{2},$$

同理可求得

$$\langle\mathbf{v},\mathbf{u}_2\rangle=\langle\mathbf{v},\mathbf{u}_3\rangle=\frac{1}{2}.$$

因此

$$\begin{aligned}\mathbf{w}&=(1,0,0,0)-\frac{1}{2}\mathbf{u}_1-\frac{1}{2}\mathbf{u}_2-\frac{1}{2}\mathbf{u}_3\\&=(1,0,0,0)-\frac{1}{4}(1,1,1,1)-\frac{1}{4}(1,-1,1,-1)\\&\quad-\frac{1}{4}(1,1,-1,-1)\\&=\frac{1}{4}(1,-1,-1,1).\end{aligned}$$

最后,正规化 $\mathbf{w}$,我们得到

$$\mathbf{u}_4=\frac{1}{2}(1,-1,-1,1),$$

$\mathbf{u}_4$ 和 $\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3$ 构成 $\mathscr{E}^4$的一组正规正交基.

我们看到这样取 $\mathbf{v}$ 是合适的. 如果错误地把 $\mathbf{v}$ 取作了 $\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3$ 的一个线性组令,那末相应的 $\mathbf{w}$ 变为一个零向量.

在无限维欧几里得向量空间中,可利用格拉姆-施密特正交化过程生成一系列正交集合. 给了我们一个有限正交集时，总可能找一个向量正交于所给集合中的每一个向量. 我们下面举出多项式方面的一个例子. 我们应该提到,在实际问题中,常常用与实际

所考虑的问题密切联系的方法，来代替生成正交集的格拉姆-施密特正交化过程．我们将在 5.5 节中讨论这一点．

例 17 前三个勒让德多项式是

$$P_0(x)=1,\quad P_1(x)=x,\quad P_2(x)=\frac{1}{2}(3x^2-1).$$

对内积

$$\langle f,g\rangle=\int_{-1}^{1}f(x)g(x)dx,$$

P_0,P_1 和 P_2 构成一个正交集(但不是正规正交的)．

求一个三次多项式 $P_3(x)$，使 P_0,P_1,P_2,P_3 构成一个正交集．

为了利用公式(17)，先正规化 $P_i(i=0,1,2)$．函数

$$f_0=\sqrt{\frac{1}{2}}P_0,\quad f_1=\sqrt{\frac{3}{2}}P_1,\quad f_2=\sqrt{\frac{5}{2}}P_2$$

有单位长度．设 $f(x)=x^3$．显然 f 不是 f_0,f_1,f_2 的线性组合．根据公式(17)，函数

$$g=f-\sum_{i=0}^{2}\langle f,f_i\rangle f_i$$

正交于 f_0,f_1 和 f_2，于是也正交于 P_0,P_1 和P_2．我们先求出

$$\langle f,\ f_0\rangle=\int_{-1}^{1}x^3\sqrt{\frac{1}{2}}dx=0$$

$$\langle f,\ f_1\rangle=\int_{-1}^{1}x^3\sqrt{\frac{3}{2}}x\,dx=\frac{2}{5}\sqrt{\frac{3}{2}}$$

$$\langle f,\ f_2\rangle=\int_{-1}^{1}x^3\sqrt{\frac{5}{2}}\cdot\frac{1}{2}(3x^2-1)dx=0.$$

因此

$$g(x)=x^3-\frac{2}{5}\cdot\sqrt{\frac{3}{2}}\sqrt{\frac{3}{2}}P_1(x)=x^3-\frac{3}{5}x.$$

定义勒让德多项式 $P_3(x)$ 为

$$P^3(x)=\frac{5}{2}g(x)=\frac{1}{2}(5\,x^3-3\,x).$$

这样调整常数是为了使 $P_3(1)=1$。正如所需，P_3 正交于 P_0, P_1 和 P_2。

练　习

22．求一个向量 $\mathbf{v}_3$，使 $\mathbf{v}_3$ 和 $\mathbf{v}_1=(2,3,-1)$，$\mathbf{v}_2=(1,1,5)$ 构成一个生成 $\mathscr{E}^3$ 的正交集。

23．求 $\mathscr{E}^3$ 的一组正规正交基，使它包含正交的单位向量

$$\mathbf{u}_1=\left(0,\sqrt{\frac{1}{2}},-\sqrt{\frac{1}{2}}\right),\quad \mathbf{u}^2=\left(\sqrt{\frac{1}{2}},\frac{1}{2},\frac{1}{2}\right).$$

24．在 $\mathscr{E}^3$ 中求一个单位向量，使它正交于

$$\mathbf{v}_1=(1,1,1)\ \text{和}\ \mathbf{v}_2=(2,0,-3).$$

（本题不能用公式(17)，因为 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 不正交。但可用类似于格拉姆-施密特正交化过程的方法来解这个问题。）

25．向量 $f_1,\cdots,f_n$ 构成欧几里得向量空间 $\mathscr{E}$ 的一组正规正交基。证明 $\mathscr{E}$ 中任何一个向量 f 可表示为如下形式

$$f=c_1f_1+\cdots+c_nf_n,\quad \text{这里}\quad c_k=\langle f,\ f_k\rangle\quad (k=1,\cdots,n).$$

这些纯量 $c_1,\cdots,c_n$ 称为 f 关于基 $f_1,\cdots,f_n$ 的分量。

26．在 $\mathscr{E}^3$ 中，对基

$$\mathbf{u}_1=\left(\frac{1}{3},\ \frac{2}{3},-\frac{2}{3}\right),\quad \mathbf{u}_2=\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right),\quad \mathbf{u}_3=\left(\frac{2}{3},-\frac{2}{3},-\frac{1}{3}\right)$$

求出 $\mathbf{v}=(1,1,1)$ 的分量。

27．证明例 17 的多项式 P_0,P_1,P_2 和 P_3 构成次数 $\leqslant 3$ 的所有多项式向量空间的一组基。对这一组基，求多项式 $P(x)=x^3+x^2+1$ 的分量。

28．证明在 $\mathscr{F}_{-1,1}$ 中，不存在次数 $\leqslant 3$ 的非平凡多项式正交于勒让德多项式 P_0,P_1,P_2 和 P_3。再证明 P_3 正交于每个次数 $\leqslant 2$ 的多项式。

29．对内积

$$\langle f, g\rangle=\int_0^{\infty} e^{-x}f(x)g(x)dx,$$

求一个三次多项式 L_3，使它和拉盖尔多项式 L_0, L_1, L_2(5.2 节练习 9)一起构成一个正交集．把多项式 $P(x)=x^3+x^2+1$ 表示为这些多项式的线性组合．

30．给出 $\mathscr{E}^3$ 中的正交单位向量 $\mathbf{u}_1, \mathbf{u}_2$，以及一个任意向量 $\mathbf{v}$，求形如 $\mathbf{v}-c_1\mathbf{u}_1-c_2\mathbf{u}_2$ 的最短向量．

[提示：见 4.2 节练习 9，并研究图 33．]

31．证明练习 30 的推广：设 $f_1, \cdots, f_m$ 是欧几里得向量空间 $\mathscr{E}$ 的一组正规正交集，并设 f 是 $\mathscr{E}$ 中任意一个向量，则向量

$$g=f-\sum_{k=1}^{m}\langle f, f_k\rangle f_k$$

在所有形如 $f-c_1f_1-\cdots-c_mf_m$ 的向量中是最短的．这个向量正交于每个 f_k $(k=1, \cdots, m)$．特别，如果 f 是 $f_1, \cdots, f_m$ 的线性组合，则 $g=0$．在 $f_1, \cdots, f_m$ 的线性组合中，$\sum_{k=1}^{m}\langle f, f_k\rangle f_k$ 是向量 f 的最好近似（对于 $\mathscr{E}$ 中的模（长度））．

32．设 $f_1, \cdots, f_m$ 是欧几里得向量空间 $\mathscr{E}$ 的一个正规正交集．又设 f 是 $\mathscr{E}$ 中一个向量，并且 $c_k=\langle f, f_k\rangle (k=1, \cdots, m)$．证明贝塞耳不等式

$$\sum_{k=1}^{m}c_k^2\leqslant \| f\|^2,$$

等式当且仅当 f 是 $f_1, \cdots, f_m$ 的线性组合时才成立．

[提示：在练习 31 中令 $\langle g, g\rangle=\| g\|^2\geqslant 0$．]

33．在 $\mathscr{F}_{-\pi, \pi}$ 中，三角函数

$$C_0(x)=\frac{1}{\sqrt{2\pi}},\ \Big\{C_k(x)=\frac{1}{\sqrt{\pi}}\cos kx,\ S_k(x)=\frac{1}{\sqrt{\pi}}\sin kx\Big\}(k=1, 2, \cdots, m)$$

构成一个正规正交集(5.2 节例 5)．设 $f(x)=x$．证明对一切 k

$$\langle f, S_k\rangle=\int_{-\pi}^{\pi} x\frac{\sin kx}{\sqrt{\pi}}dx=(-1)^{k+1}\frac{2\sqrt{\pi}}{k}$$

且

$$\langle f, C_k\rangle=0.$$

再证明对于$\mathscr{T}_{-\pi,\pi}$的模，在上述三角函数的线性组合中

$$2\left(\sin x-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\cdots+(-1)^{m+1}\frac{\sin mx}{m}\right)$$

给出了$f(x)(-\pi\leqslant x\leqslant\pi)$的最好近似。无穷级数

$$2\sum_{k=1}^{\infty}(-1)^{k+1}\frac{\sin kx}{k}$$

称为f在区间$-\pi\leqslant x\leqslant\pi$上的傅里叶级数。在开区间

$$-\pi<x<\pi$$

的所有点上这傅里叶级数收敛于f。然而在端点，$f(\pm\pi)=\pm\pi$，而傅里叶级数取值 0。

34．证明

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots\leqslant\frac{\pi^2}{6}.$$

[提示：把贝塞尔不等式用于练习 33，就是该不等式。]

有关傅里叶级数的理论可查阅 E. Kreyszig, *Advanced Engineering Mathematics*, Second Edition, Wiley, 1967.

5.5 微分方程的正交解

在物理和化学理论中出现的许多函数——所谓特殊函数，定义为某些微分方程的解。其中某些函数对于特定的内积适合正交关系。我们将说明怎样可从有关的微分方程直接证明正交性。

例 18 勒让德方程

设n是一个常数。微分方程

$$(1-x^2)y''-2xy'+n(n+1)y=0 \tag{18}$$

称为勒让德方程。$\left(\text{我们分别用标准记号 } y' \text{ 和 } y'' \text{ 代替 } \frac{dy}{dx} \text{ 和 } \frac{d^2y}{dx^2}.\right)$

当n是一个非负整数时，勒让德方程有n次多项式的解，所有这些多项式解都是一个特定多项式解的纯量倍。它还具有非多

项式的解，但在这里我们不关心这些解．n次多项式解$y=P_n(x)$如果使$P_n(1)=1$，则称它为n次勒让德多项式．我们现在不解方程(18)而证明，若$n\neq m$，则对于内积

$$\langle f,g\rangle=\int_{-1}^{1}f(x)g(x)dx,$$

P_n正交于P_m．

我们已知

$$(1-x^2)P_n''-2xP_n'+n(n+1)P_n=0 \tag{19}$$

与

$$(1-x^2)P_m''-2xP_m'+m(m+1)P_m=0. \tag{20}$$

把方程(19)乘以P_m，把方程(20)乘以P_n，然后两式相减，我们得到

$$(1-x^2)(P_n''P_m-P_m''P_n)-2x(P_n'P_m-P_m'P_n)$$
$$=\{m(m+1)-n(n+1)\}P_nP_m.$$

这个方程化简为

$$\frac{d}{dx}\{(1-x^2)(P_n'P_m-P_m'P_n)\}=(m-n)(m+n+1)P_nP_m.$$

再在区间$-1\leqslant x\leqslant 1$上积分，我们得到

$$[(1-x^2)(P_n'(x)P_m(x)-P_m'(x)P_n(x)]_{-1}^{1}$$
$$=(m-n)(m+n+1)\int_{-1}^{1}P_n(x)P_m(x)dx.$$

现在当$x=1$和$x=-1$时$(1-x^2)(P_n'(x)P_m(x)-P_m'(x)P_n(x))$总是0．因此

$$(m-n)(m+n+1)\int_{-1}^{1}P_n(x)P_m(x)dx=0.$$

而$m\neq n$，且$m\neq-(n+1)$，因为m和n都是非负整数．由此

$$\langle P_n,P_m\rangle=\int_{-1}^{1}P_n(x)P_m(x)dx=0$$

这正是所要证明的．

注意对多项式$P_n(x)$我们没有得到一个明确的表达式．我们讨论它们适合的微分方程，间接地证明了它们的正交性（当$n\neq$

m). 对照一下，在 5.4 节，我们把含有低次勒让德多项式的表达式进行积分，以此检验了它们的正交性.

注意：不要错误地以为勒让德方程任意两个解 y_n 和 y_m($n\neq m$)都是正交的. 如果你按照例 18 的论证，那末你必须证明，$(1-x^2)(y_n'(x)y_m(x)-y_m'(x)y_n(x))$在 $x=1$ 和 $x=-1$ 时都为零. 这在一般情形是不成立的，虽然它在 $y_m(x)$和 $y_n(x)$ 是多项式时显然是正确的.

例 19 埃尔米特微分方程是*

$$y''-2xy'+2ny=0, \tag{21}$$

其中 n 是一个常数. 当 n 是非负整数时，方程(21)有 n 次多项式的解. 这些多项式解都是特定解 $H_n(x)$的纯量倍，$H_n(x)$中 x^n 的系数是 2^n. $H_n(x)$称为 n 次埃尔米特多项式.

对于内积

$$\langle f,g\rangle=\int_{-\infty}^{\infty}e^{-x^2}f(x)g(x)dx, \tag{22}$$

证明当 $n\neq m$，H_n 正交于 H_m.

我们已知

$$H_n''-2xH_n'+2nH_n=0 \tag{23}$$

和

$$H_m''-2xH_m'+2mH_m=0. \tag{24}$$

用 $e^{-x^2}H_m$ 乘方程(23)并用 $e^{-x^2}H_n$ 乘方程(24)，然后两式相减，我们得到

$$e^{-x^2}\{(H_n''H_m-H_m''H_n)-2x(H_n'H_m-H_m'H_n)\}$$

---

* 在某些书中，引用的埃尔米特微分方程形如

$$y''-xy'+ny=0. \tag{21 a}$$

从方程(21)作代换 $x=z/\sqrt{2}$，并把 y 当作 z 的函数，便得到这个微分方程(21 a)。方程(21 a)的多项式解不同于(21)的多项式解。它们对于内积

$$\langle f,g\rangle=\int_{-\infty}^{\infty}e^{-z^2/2}f(z)g(z)dz$$ 是正交的。

表 4 正交函数

	微分方程	内积	正交关系
1. 多项式			
勒让德			
$P_n(x)$ $(n=0,1,2,\cdots)$	$(1-x^2)P_n''-2xP_n'+n(n+1)P_n=0$	$\int_{-1}^{1}f(x)g(x)dx$	$\langle P_n, P_m\rangle=0$ $(n\neq m)$
埃尔米特			
$H_n(x)$ $(n=0,1,2,\cdots)$	$H_n''-2xH_n'+2nH_n=0$	$\int_{-\infty}^{\infty}e^{-x^2}f(x)g(x)dx$	$\langle H_n, H_m\rangle=0$ $(n\neq m)$
拉盖尔			
$L_n(x)$ $(n=0,1,2,\cdots)$	$xL_n''+(1-x)L_n'+nL_n=0$	$\int_{0}^{\infty}e^{-x}f(x)g(x)dx$	$\langle L_n, L_m\rangle=0$ $(n\neq m)$
车贝雪夫			
$T_n(x)$ $(n=0,1,2,\cdots)$	$(1-x^2)T_n''-xT_n'+n^2T_n=0$	$\int_{-1}^{1}\frac{1}{\sqrt{(1-x^2)}}f(x)g(x)dx$	$\langle T_n, T_m\rangle=0$ $(n\neq m)$
2. 三角函数			
(正规化的)			
$S_n(x)=\frac{1}{\sqrt{\pi}}\sin nx$ $(n=1,2,\cdots)$	$S_n''+n^2S_n=0$	$\int_{-\pi}^{\pi}f(x)g(x)dx$	$\langle S_n, C_m\rangle=0$ (一切 m,n)
$C_n(x)=\frac{1}{\sqrt{\pi}}\cos nx$	$C_n''+n^2C_n=0$		$\langle S_n, S_m\rangle=0$ $(n\neq m)$
$C_0(x)=\frac{1}{\sqrt{2\pi}}$			$\langle C_n, C_m\rangle=0$

$$=2(m-n)e^{-x^2}H_nH_m.$$

再在区间$-\infty<x<\infty$上积分，我们得到

$$[e^{-x^2}(H_n'(x)H_m(x)-H_m'(x)H_n(x))]_{-\infty}^{\infty}$$

$$=2(m-n)\int_{-\infty}^{\infty}e^{-x^2}H_n(x)H_m(x)dx.$$

现在 $H_n'(x)H_m(x)-H_m'(x)H_n(x)$是一个多项式，且不难证明对一切多项式 $P(x)$有

$$\lim_{x\to\pm\infty}e^{-x^2}P(x)=0.$$

因此

$$2(m-n)\int_{-\infty}^{\infty}e^{-x^2}H_n(x)H_m(x)dx=0.$$

因为 $m\neq n$,所以得知对于内积(22),H_n 和 H_m 是正交的.

在上页表 4 中，我们列举了在理论物理和理论化学中出现的一些函数,以及它们所适合的正交关系.

练　习

35. 验证拉盖尔和车贝雪夫微分方程的多项式解适合表 4 中指明的正交关系。

36. 验证 5.4 节中见到的勒让德多项式适合勒让德微分方程。类似地考虑埃尔米特和拉盖尔多项式.

37. (i) 证明，如果 $P_n(x)$是次数 $n\geqslant 1$ 的适合勒让德方程的一个多项式，则

$$\int_{-1}^{1}P_n(x)dx=0.$$

[提示：利用正交关系.]

(ii) 对埃尔米特，拉盖尔，车贝雪夫多项式，叙述和证明类似的结果.

有关勒让德多项式，埃尔米特多项式等等的理论和应用的论述,可参阅 I. N. Sneddon, *Special Functions of Mathematical Physics and Chemistry*, Oliver & Boyd, 1956.

5.6 酉空间

这一章到此为止，我们考虑的是实数上的向量空间．我们已经看到，在讨论物理和化学的特殊函数时，内积是一种很有用的工具．许多物理理论是基于复函数的（量子理论便是一个例子），而且在那里，内积也是一种重要工具．然而，为了保持长度的正定性，必须修改内积的定义．对于实函数，这新定义与原定义一致．

我们用 $\bar{z}$ 表示复数 z 的复共轭．于是，如果 $z=a+ib$（a, b 为实数），则 $\bar{z}=a-ib$．我们看到 $\bar{\bar{z}}=z$ 与 $|z|^2=\bar{z}z=a^2+b^2$．除了 $a=b=0$ 的情况之外，$|z|^2$ 总是一个正实数；当 $a=b=0$ 时，$|z|^2=0$．此外，$z=\bar{z}$ 当且仅当 z 为实数．

设 $\mathscr{V}$ 是复数上的向量空间．如果对任意一对向量 f,g，定义复数 $\langle f,g\rangle$ 使得

(i) $\langle f,g\rangle=\overline{\langle g,f\rangle}$　（共轭对称性）；

(ii) 对一切向量 f,g,h 及纯量（复数）k,l，

$$\langle h,kf+lg\rangle=\bar{k}\langle h,f\rangle+\bar{l}\langle h,g\rangle \qquad \text{（共轭双曲性）}$$

及

$$\langle kf+lg,h\rangle=k\langle f,h\rangle+l\langle g,h\rangle;$$

(iii) 如果 $f\neq\theta$，则 $\langle f,f\rangle>0$；而 $\langle\theta,\theta\rangle=0$（正定性），

则函数 $\langle\,,\,\rangle$ 称为 $\mathscr{V}$ 上的内积，而 $\mathscr{V}$ 称为酉向量空间．

(ii) 中第二个关系式是头一个关系式和共轭对称性的推论．

仅从共轭对称性，便得出 $\langle f,f\rangle$ 是实数，因为 $\langle f,f\rangle=\overline{\langle f,f\rangle}$．在(iii)中我们还要求 $\langle f,f\rangle$ 是正定的．这样，就使我们可以把长度概念合理地推广：在酉空间中，向量 f 的长度是

$$\|f\|=\sqrt{\langle f,f\rangle}.$$

我们定义，酉空间中两个向量 f 和 g 是正交的，当且仅当 $\langle f,g\rangle=0$．这就建立了正交性的概念．这个定义关于 f 和 g 是对称的，因为 $\langle f,g\rangle=0$ 当且仅当 $\langle g,f\rangle=0$．

例 20 两个重要的酉向量空间

(i) 设 $\mathscr{C}^n$ 是 1.5 节所定义的由所有带复分量的向量 $\mathbf{v}=(v_1,\cdots,v_n)$，组成的向量空间. 定义内积为

$$\langle \mathbf{v},\mathbf{w}\rangle=\bar{v}_1w_1+\bar{v}_2w_2+\cdots+\bar{v}_nw_n. \tag{25}$$

共轭对称性，共轭双线性和正定性都是满足的.

于是赋于了内积 (25) 的 $\mathscr{C}^n$ 是一个酉空间，我们用 $\mathscr{U}^n$ 记之.

(ii) 所有在实区间 $a\leqslant x\leqslant b$ 连续的复值函数，定义内积

$$\langle f,g\rangle=\int_a^b\overline{f(x)}g(x)dx, \tag{26}$$

这样的无限维空间是一个酉空间.

对欧几里得空间建立的大部分理论都可推广到酉空间上去，但它们的证明可能需要稍微修改一下.

例 21 在 $\mathscr{U}^3$ 中，求 $\mathbf{v}=(1+i,i,1)$ 的长度，并证明 $\mathbf{v}$ 正交于 $\mathbf{w}=(1-i,-1,i)$.

$$\begin{aligned}\|\mathbf{v}\|^2=\langle\mathbf{v},\mathbf{v}\rangle&=(\overline{1+i})(1+i)+\bar{i}i+(\bar{1})1\\&=(1-i)(1+i)+(-i)i+1=4.\end{aligned}$$

因此 $\|\mathbf{v}\|=2$.

还有，

$$\begin{aligned}\langle\mathbf{v},\mathbf{w}\rangle&=(\overline{1+i})(1-i)+\bar{i}(-1)+\bar{1}(i)\\&=(1-i)(1-i)+i+i=0,\end{aligned}$$

因此 $\mathbf{v}$ 和 $\mathbf{w}$ 是正交的.

例 22 $f_1,\cdots,f_m$ 是一个酉空间中相互正交的向量. 证明 $f_1,\cdots,f_m$ 线性无关.

其证明与欧几里得空间情况下的证明(5.4 节例 13)类似.

例 23 构造正规正交基

在 $\mathscr{U}^3$ 中，求包含正交单位向量 $\mathbf{u}_1=\frac{1}{2}(1+i,\ i,\ 1)$ 和 $\mathbf{u}_2=$

$\frac{1}{2}(1-i,-1,i)$ 的一组正规正交基.

只要把格拉姆-施密特正交化过程稍加修改便可应用. 取向量 $\mathbf{v}$,使它不是 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 的线性组合,并作

$$\mathbf{w}=\mathbf{v}-\sum_{i=1}^{2}\overline{\langle\mathbf{v},\mathbf{u}_i\rangle}\mathbf{u}_i, \tag{27}$$

则 $\mathbf{w}$ 和 $\mathbf{u}_1,\mathbf{u}_2$ 都正交.

选取 $\mathbf{v}$ 时必须小心. 向量(1,0,0)是不适用的,因为

$$(1,0,0)=\frac{1}{2}(1-i)\mathbf{u}_1+\frac{1}{2}(1+i)\mathbf{u}_2,$$

它是 $\mathbf{u}_1,\mathbf{u}_2$ 的线性组合. 然而,我们可取 $\mathbf{v}=(0,1,0)$,则

$$\langle\mathbf{v},\mathbf{u}_1\rangle=\frac{1}{2}i,\ \langle\mathbf{v},\mathbf{u}_2\rangle=-\frac{1}{2}\ \text{和}\ \overline{\langle\mathbf{v},\mathbf{u}_1\rangle}=-\frac{1}{2}i,\ \overline{\langle\mathbf{v},\mathbf{u}_2\rangle}=-\frac{1}{2}.$$

因此 $\mathbf{w}=(0,1,0)+\frac{1}{4}i(1+i,i,1)+\frac{1}{4}(1-i,-1,i)$

$$=\left(0,\frac{1}{2},\ \frac{1}{2}i\right).$$

再正规化,我们得到

$$\mathbf{u}_3=\sqrt{\frac{1}{2}}(0,1,i),$$

它是正交于 $\mathbf{u}_1$ 和 $\mathbf{u}_2$ 的单位向量.

怎样在酉空间中用格拉姆-施密特正交化过程,以及任何有限生成酉空间有正规正交基,这两点现在应当清楚了.

我们用一个实变量复函数的例子作为结束.

例 24 定义函数

$$f_n(x)=\frac{1}{\sqrt{2\pi}}e^{inx}=\frac{\cos nx+i\sin nx}{\sqrt{2\pi}}\ (n=0,\pm1,\pm2,\cdots),$$

对内积

$$\langle f,g\rangle=\int_{-\pi}^{\pi}\overline{f(x)}g(x)dx,$$

证明 f_n 构成一个正规正交集.

设 m 和 n 是不同的整数，则

$$\langle f_n,f_m\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-inx}e^{imx}dx=\frac{1}{2\pi i(m-n)}[e^{i(m-n)x}]_{-\pi}^{\pi}$$

$$=\frac{1}{2\pi i(m-n)}(e^{i(m-n)\pi}-e^{-i(m-n)\pi})=0,$$

最后一步是因为当 k 是整数时 $e^{ik\pi}=e^{-ik\pi}=(-1)^k$. 由此可见这些 f_n 组成正交集. 而 f_n 的长度是

$$\|f_n\|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-inx}e^{inx}dx=1.$$

结论证毕.

注意，函数 f_n 是微分方程

$$\frac{d^2y}{dx^2}+n^2y=0$$

的解.

酉空间是量子力学理论的基础. 例如，可参阅 R. A. Newing 和 J. Cunningham, *Quantum Mechanics*, Oliver & Boyd, 1967.

练　习

38. 证明$\left(\frac{1}{2},\frac{1}{2}i,\sqrt{\frac{1}{2}}i\right)$，$\left(\frac{1}{2}i,-\frac{1}{2},\sqrt{\frac{1}{2}}\right)$，$\left(\frac{1}{2}+\frac{1}{2}i,\frac{1}{2}-\frac{1}{2}i,0\right)$构成 $\mathscr{U}^3$ 的一组正规正交基.

39. 求两个向量，使它们和 $\mathbf{u}_1=\frac{1}{2}(1,i,1,i)$，$\mathbf{u}_2=\frac{1}{2}(i,1,i,1)$一起构成 $\mathscr{U}^4$ 的一组正规正交基.

40. 证明第二个共轭双线性关系是头一个共轭双线性关系和共轭对称

性的推论。

41. 在$\mathscr{U}^2$中设$\mathbf{v}=(1,1)$，$\mathbf{w}=(i,i)$。证明

$$\|\mathbf{v}+\mathbf{w}\|^2=\|\mathbf{v}\|^2+\|\mathbf{w}\|^2.$$

再证明$\langle\mathbf{v},\mathbf{w}\rangle\neq 0$。

42. 对酉空间证明毕达哥拉斯定理（5.2 节例 7）。练习 41 说明这个定理的逆对酉空间是不成立的。问题出在证明的哪一步？

43. 在酉空间中，证明对任意复纯量k有

$$\|kf\|=|k|\|f\|.$$

44. 在酉空间中，证明

$$\|f-kg\|^2=\|f\|^2-2\,\mathrm{Re}(k\langle f,g\rangle)+|k|^2\|g\|^2,$$

这里$\mathrm{Re}(z)$表示复数z的实部。

45. 证明许瓦尔兹不等式在酉空间成立。

46. 函数f和g是实变量x的复值函数。证明

$$\left|\int_0^1\overline{f(x)}g(x)dx\right|^2\leqslant\int_0^1|f(x)|^2dx\int_0^1|g(x)|^2dx.$$

第二部分

对称变换群

第六章　对称变换

6.1　引论

我们都熟悉对称这个概念．比如，在糊墙纸的图案里，我们可以看到对称；大家认为字母 **A** 与 **O** 比字母 **P** 更为对称．物理学家也重视对称．举例来说，我们来看一个氨分子 NH_3．在这个分子里，三个氢原子位于一个等边三角形的三个顶点上，而氮原子与三个氢原子的距离相等．完全有理由猜测，氢原子的这种对称排列必然会以某种方式在氨分子的物理性质中体现出来．实际上也确实如此，氨分子的电子分布以及振动光谱就与氢原子的对称排列有关．又如在研究晶体性质时，晶体内部结构的对称性起着重要作用．

以上我们仅是粗略地介绍了对称这个概念．为了深入讨论对称，我们必须根据条件找到一套方法，来刻划事物的对称特性．现在我们就来找这样的方法．

例 1　一个等边三角形的对称性

为简单起见，我们先考虑等边三角形这样一个二维例子（图 34）．假如我们将此三角形绕它的中心 O 逆时针旋转 $120°$ 角 $\left(=\frac{2\pi}{3}\text{弧度}\right)$，那末此三角形经旋转后的位置与原来的位置完全重合．这是为什么？其原因就是这个等边三角形的对称性！又假如我们将这个三角形所在平面的点关于直线 L 作反射*，那末反射前后两个三角形位置完全重合．

---

* 在二维问题中，反射"镜面"是一条直线．关于直线 L 作反射后，P 点的"象"是 P' 点，这里直线 L 垂直平分 PP'．直线 L 两边的点都要进行反射．

于是我们有一个度量对称性的方法：列举出使该图形不变形地变到与自身重合的平面变换．这种变换称为等边三角形的对称变换．三角形的三个顶点应该看成是无区别的，但为了讨论时便于枚举出对称变换，我们将三个顶点标上字母．在下面的表 5 中，我们列出了该等边三角形的六个不同的对称变换．

还有其它的使三角形变到其自身的方法，例如，将三角形逆时针旋转 $2\pi+\frac{2}{3}\pi=\frac{8}{3}\pi$．但是，所得到的(标字母的)图形与逆时针旋转$\frac{2}{3}\pi$所得的图形是相同的．所以，若二个对称变换在图形上作用的方式一致，则我们将这二个对称变换看成是同一个变换．

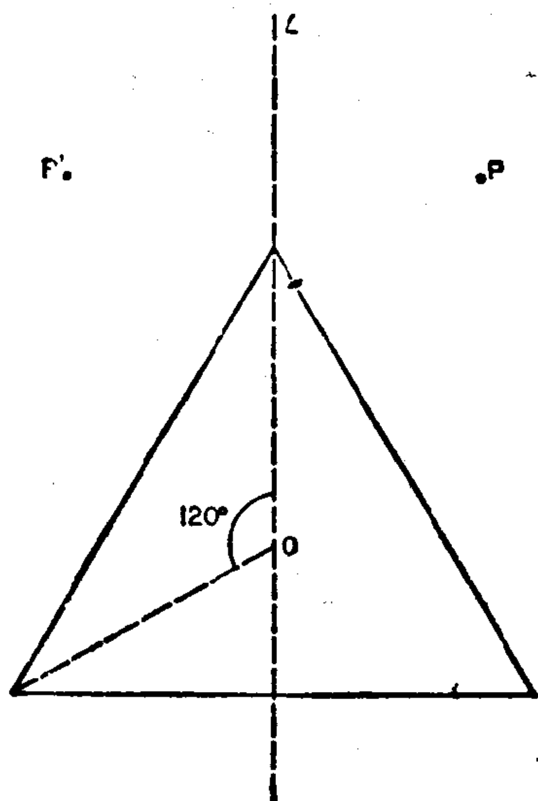


图 34 一个等边三角形的对称性

由于表 5 最终位置栏里的六个三角形的顶点位置互不相同，所以表中的六个对称变换是不同的对称变换．而且，表 5 列出了等边三角形的所有不同的对称变换．为了证明这一点，只须注意在任意一个对称变换下，顶点 a 必须变到三个位置(顶上、左边或右边)之一．在每种情况下，顶点 b 及顶点 c 的位置只有两种可能的配置情况．因此，不可能有多于$3\times2=6$个不同的对称变换．可见恰有六个不同的对称变换．

表 5 中的最后一个对称变换“保持原位”叫做恒等变换．在枚

表 5

初始位置（顶点标上字母）	对称变换	最终位置（顶点标上字母）
L_1 L_2 L_3 a b c	1．绕中心逆时针旋转$\frac{2}{3}\pi$	c a b
	2．绕中心逆时针旋转$\frac{4}{3}\pi$	b c a
	3．关于 L_1 的反射	a c b
	4．关于 L_2 的反射	c b a
	5．关于 L_3 的反射	b a c
	6．保持原位置	a b c

举对称变换时，把这个平凡变换包括进去是方便的．如果一个图形仅有的对称变换是恒等变换，则我们认为这个图形是非常不对称的．其它的图形都有一些非平凡的对称变换，即具有某种程度的对称性．

例 2 讨论字母 S，O，A 以及 P 的对称性．

P 字母P是非常不对称的．另外的三个字母都有非平凡的对称变换，如表 6 所示．

表 6

字母	对称变换
A	1．恒等变换 2．关于 L 的反射
S	1．恒等变换 2．关于中心的 π 角旋转
O	1．恒等变换 2．关于中心的任意角旋转 3．关于过中心的任意直线的反射

字母 A 与 S 恰有二个不同的对称变换，字母 O 有无数多个对称变换．

例 3 正方形的对称性

表 7 中列出了八个不同的对称变换．还有别的对称变换吗？答案是没有．因为为了使正方形不变形，顶点 a 及顶点 c 必须处于对角位置，共有四种这样的可能情况．对顶点 a 及顶点 c 的每个可能位置，相应顶点 b 及顶点 d 的位置只有二种可能．所以，正方形对称变换不可能多于 $4\times2=8$ 个．

表 7　正方形的对称变换

初始位置(顶点标上字母)	对称变换	最终位置(顶点标上字母)
L_1 L_3 L_2 L_4 b a c d	1. 恒等变换	b a c d
	2. 逆时针旋转$\frac{\pi}{2}$	a d b c
	3. 逆时针旋转π	d c a b
	4. 逆时针旋转$\frac{3}{2}\pi$	c b d a
	5. 关于 L_1 的反射	a b d c
	6. 关于 L_2 的反射	c d b a
	7. 关于 L_3 的反射	d a c b
	8. 关于 L_4 的反射	b c a d

现在我们来看三维的例子.

例 4 氨(NH_3)分子的对称性(图35(a)与(b)).

三个氢原子分别在平面 π_1,π_2 及 π_3 上,它们是相同的三个氢原子.

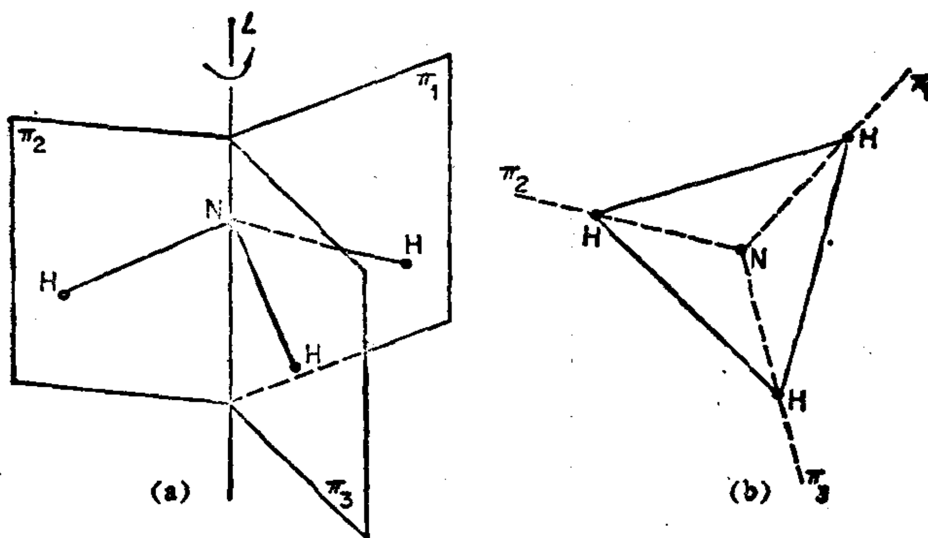


图 35 氨分子的对称性

若一个对称变换使氨分子与自身重合,那末这个对称变换必然使氮原子固定不动.这个分子的平面草图(图35(b))使我们想起例1,从而我们可列举出该分子的下列对称变换:

1. 恒等变换.
2. 绕轴 L(顺箭头所指方向)旋转 $\frac{2}{3}\pi$.
3. 绕轴 L 旋转 $\frac{4}{3}\pi$.
4. 关于平面 π_1 的反射.
5. 关于平面 π_2 的反射.
6. 关于平面 π_3 的反射.

氨分子只有以上六种不同的对称变换.其证明与例1的证明一样.

最后,我们来看一个晶体学中的例子.

例 5 石墨由许多碳原子的平面六边形薄片构成(图36). 在考虑石墨内部的微小结构时，这些薄片可以看成是无限延伸的. 试讨论其中一片的对称性.(在蜂房及一些浴室的地面上也有这样的六边形图案.)

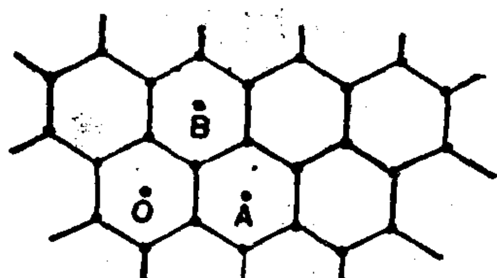


图 36 石墨薄片的对称性

在枚举本例的对称变换时，我们要遇到平移这种对称变换. 将每个点沿着某固定方向移动一个固定距离的变换称为平移. 我们可用一个向量来代表一个平移，该向量的方向就是平移的方向，向量的大小等于平移所移动的距离的大小.

在此薄片所在的平面上，固定相邻三个六边形的中心 O、A 及 B. 令 $OB=AB=d$. 我们可以枚举这个图案的对称变换如下：

1. 将图案沿向量 $\overrightarrow{OA}$ 的方向(即向右)或沿向量 $\overrightarrow{AO}$ 的方向(即向左)移动距离 d.

2. 沿 $\overrightarrow{OB}$ 或 $\overrightarrow{BO}$ 的方向移动距离 d.

3. 绕 O 点逆时针旋转 $\frac{2}{6}\pi$.

4. 以过 O 点及 A 点的直线作反射.

5. 任意连续实施上面四种类型的变换.

上面第五种变换是对称变换，这是因为依次实施一个图形的二个(或更多)对称变换仍是这个图形的对称变换. 在枚举一个图形的对称变换时，记住这一点是很重要的.

可以证明可用以上五种类型的对称变换来表示此种六边形图案的任意对称变换. 学了变换的代数后，就很容易证明这一点*.

---

* 见(8.1)节.

下一节里，我们讨论变换的代数。

练　习

1. 讨论以下图形的对称性：

(i) 正五边形。

(ii) 正六边形。

(iii) 菱形。

(iv) 邻边不相等的平行四边形。

(v) 椭圆。

(vi) 抛物线。

2. 证明可用相同的对称变换来刻划字母 H 的对称性以及邻边不相等的平行四边形的对称性。

3. 将所有大写字母分组，使每组中的字母有相同的对称性。

4. 证明平移不可能是一个有限图形的对称变换。

5. 研究糊墙纸及瓷砖地面等物体的对称性。

6.2 变换的代数

使平面上每个点 P 对应唯一的点 P' 的一个规则称为平面的一个变换（或映射），记成 T。点 P' 称为点 P 在变换 T 下的象，可用函数的形式写成

$$P'=T(P).$$

设变换 S 及变换 T 是同一平面的二个变换．若对该平面上的所有点 P，$S(P)=T(P)$，则称这二个变换相等，记为 $S=T$．例如，绕 O 点逆时针旋转 $\frac{\pi}{2}$ 及 $\frac{5}{2}\pi$ 是二个相等的变换．有许多变换使二个不同的点有不同的象，例如，旋转、反射及平移这些变换都具有这种性质，这样的变换称为一一变换．

平面的变换不全是一一变换．例如，到直线 L 上的正交射影（图37）．在过 P_1 及 P_2 二点直线上的所有点有相同的象 P'．正交射影是一个多一变换，即平面上许多不同的点有相同的象．

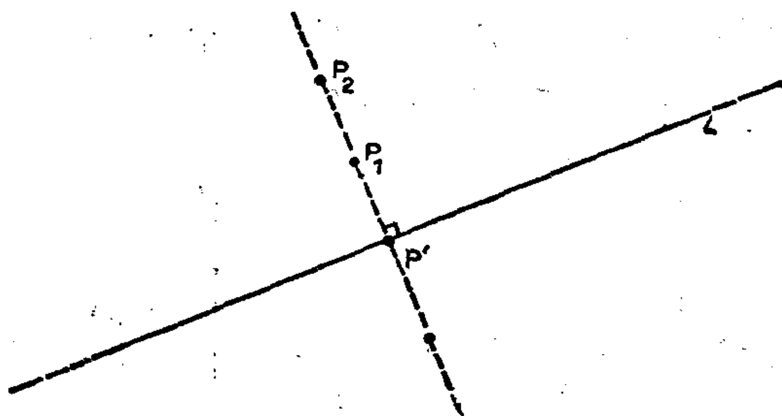


图 37 正交射影

变换的乘积

设 S 与 T 是同一平面的二个变换．我们定义 S 与 T 的乘积 ST 是一个变换，它对平面上的所有点 P，下式成立：

$$ST(P)=S(T(P)) \tag{1}$$

称 ST 是由 S 及 T 相乘得出的．

设

$$T(P)=P' \quad 及 \quad S(P')=P''.$$

则

$$ST(P)=P''.$$

所以，点 P 在变换 ST 下的象与点 P 先作用变换 T，然后将所得的象再经变换 S 作用后所得的象相同．但要注意二个变换作用的顺序！下图说明了这个过程：

$$P\overset{T}{\rightarrow}P'=T(P)\overset{S}{\rightarrow}P''=S(P')=S(T(P)).$$
$$\underbrace{\qquad\qquad ST \qquad\qquad}_{}$$

例 6

将以下同一平面的二个变换相乘：

R：绕 O 点逆时针旋转 $\frac{2}{3}\pi$；

S：关于过O点的直线L作反射.

由图38易见,$RS(P)$及$SR(P)$不是相同的点. 所以

$$RS \neq SR.$$

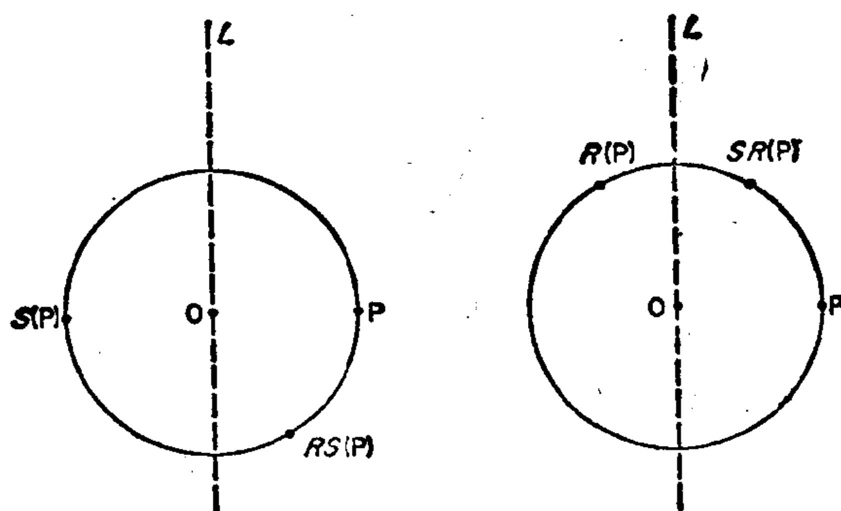


图 38　变换乘法的不可交换性

变换的乘法一般来说不是交换运算. 在求许多变换的乘积时,必须特别注意这些变换的顺序.

现在我们证明变换乘积满足结合律.

定理 I　任意三个变换R、S、T满足

$$(RS)T = R(ST). \tag{2}$$

证明　设点P是平面上任意一点,置

$$T(P)=P', \quad S(P')=P'', \quad R(P'')=P'''.$$

我们将要证明$(RS)T$及$R(ST)$这二个变换均将点P映到点P'''. 为简便,令$RS=W$,得

$$\begin{aligned}(WT)(P) &= W(T(P)) \ (\text{定义}) \\ &= W(P') = (RS)(P') \\ &= R(S(P')) \ (\text{定义}) \\ &= R(P'') = P'''.\end{aligned}$$

对$R(ST)$实施相同的论证,就能完成本定理的证明.

作用变换$(RS)T$或作用变换$R(ST)$所得的结果，都与先作用T，然后作用S，最后作用R所得的结果相同. 但要注意作用的顺序！所以，在写变换的乘积时，省略掉其中的括号是不会引起误解的. 例如，

$$(RS)T=R(ST)=RST.$$

可用归纳法证明n个变换，只要它们相乘的次序决定了，则无论怎样加括号，所得的乘积是相同的. 例如，

$$(T_1T_2)(T_3T_4)=(T_1(T_2T_3))T_4=T_1T_2T_3T_4,$$

它们都等于依次实施变换T_4,T_3,T_2,T_1. 我们已看到有些变换相乘时，不可交换顺序，因此必须十分注意变换相乘时的次序.

恒等变换

恒等变换1是一个平凡的变换，但又是一个重要的变换. 恒等变换的定义如下：对所有点P，使得

$$1(P)=P.$$

恒等变换相当于"使所有的点都不动"的规则。

恒等变换与任意变换T可以变换，显然有

$$1\,T=T\,1=T. \tag{3}$$

所以用1乘变换与用数1乘数十分相似。

变换的逆

如图39，T是绕O点逆时针旋转α角的变换，S是绕O点顺时针旋转α角的变换. 变换T与变换S互相抵消，所以变换ST使每个点不动，变换TS也使每个点不动. 得

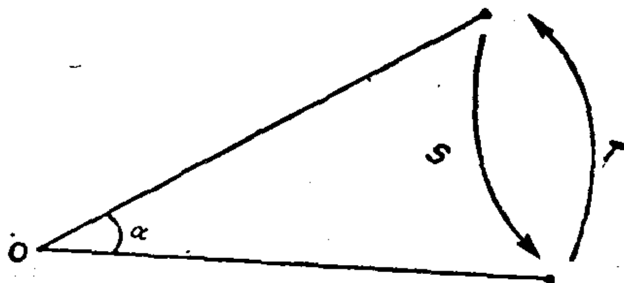


图39 变换与逆

$$ST=TS=\mathbf{1}. \tag{4}$$

一般地，如果S与T是满足(4)式的二个变换，则称S是T的逆(或逆变换)，记成

$$S=T^{-1}. \tag{5}$$

因为(4)式是对称的，所以从S是T的逆，可推出T也是S的逆．由此可推得

$$T=S^{-1}=(T^{-1})^{-1}. \tag{6}$$

不是所有的变换都有逆，例如，正交射影就没有逆．

容易证明：如果T有逆，那末其逆是唯一的．不然，我们可设S_1,S_2是T的二个不同的逆．重复用(3)与(4)可得

$$S_1=S_1\mathbf{1}=S_1(TS_2)=(S_1T)S_2=\mathbf{1}S_2=S_2,$$

与假设矛盾．

若某个变换有逆，则称这个变换是非奇异的．没有逆的变换称为奇异的．旋转、反射以及平移都是非奇异变换的例子．

乘积的逆

设T_1,T_2是两个非奇异变换，则T_1T_2是非奇异的，而且它的逆由下式给出：

$$(T_1T_2)^{-1}=T_2^{-1}T_1^{-1}. \tag{7}$$

请注意上式中的次序！为了证明公式(7)，可作计算

$$(T_1T_2)(T_2^{-1}T_1^{-1})=T_1\mathbf{1}T_1^{-1}=\mathbf{1},$$

同理可得$(T_2^{-1}T_1^{-1})(T_1T_2)=\mathbf{1}$．所以$T_2^{-1}T_1^{-1}$是$T_1T_2$的逆．

更一般地，设$T_1, T_2, \cdots, T_n$是非奇异变换，则它们的乘积$T_1T_2\cdots T_n$是非奇异的，并且

$$(T_1T_2\cdots T_n)^{-1}=T_n^{-1}\cdots T_2^{-1}T_1^{-1}. \tag{8}$$

变换的幂

自然地将变换TT写成T^2．同样$T^3=TTT$(可省掉括号)．设n是正整数，一般地，连续实施n次变换T所得的变换记成T^n．称T^n是T的n次幂．

指数法则

$$T^mT^n=T^{m+n} \tag{9}$$

显然对正整数 m,n 成立.

若 T 是非奇异的,我们将 $(T^{-1})^n$ (n 是正整数) 记成 T^{-n}. 定义 $T^0=1$. 那末,很容易证明(9)式对所有整数 m,n 成立(整数包括正整数、负整数、零).

另一个指数法则

$$(T^m)^n=T^{mn} \tag{10}$$

对所有整数 m,n 成立. (若 T 是奇异的,则 m、n 只能取非负整数.)

例 7 设 T 是非奇异变换,求证

$$T^{-n}=(T^{-1})^n=(T^n)^{-1} \tag{11}$$

对所有整数 n 成立.

第一种情况: n 是正整数. (11)式中的第一个等式是定义. 对一切 i,令 $T_i=T$,则由(8)式可得上面的第二个等式.

第二种情况: n 是负整数. 令 $n=-m$,用公式(6),(8),(11)(对正的指数)即可.

第三种情况: $n=0$. 显然.

例 8 求证 $(T^{-5})^{-2}=T^{10}$.

$(T^{-5})^{-2}=((T^{-5})^{-1})^2$(定义)$=(T^5)^2$(公式(11))$=T^{10}$.

变换的一般理论不局限于讨论平面上的变换. 上面的讨论可推广到三维空间中的变换. 在第二章中学过的初等行变换, 可以看成是作用在矩阵的行上的变换. 这些行初等变换都是非奇异的. 作为本节最后一例,一些我们熟悉的函数,例如,正弦函数、对数函数以及指数函数也只是作用在实数上的变换而已.

练　　习

6. 设 R、S 分别是例 6 中的旋转与反射,T 是同一平面的平移并由向量

$a \neq 0$ 所定义。求证 $RT \neq TR$，S 与 T 可以交换吗？

7. 证明任意二个平移 T_1、T_2 可以交换，任意二个绕同一点 O 的平面旋转可以交换. 旋转中心不同的旋转可以交换吗？

8. 证明二个不同的反射不可以交换.

9. 补全定理 1 的证明.

10. 对 $n \geqslant 3$ 个变换，用归纳法证明结合律成立：对乘积 $T_1T_2\cdots T_n$ 无论怎样加括号，所得到的变换是相同的.

11. 设 T 是任意一个变换，求证 $1T = T1 = T$.

12. 求下列变换的逆：

(i) 由位置向量 a 定义的平移.

(ii) 关于直线 L 的反射.

13. 证明恒等变换的逆是自身. 你还知道哪些变换也具有这种性质？

14. 求证正交射影没有逆.

[提示：假定 T 有逆 S，使 $ST = 1$. 用 T 是多一映射可得出矛盾.]

15. 验证公式 (10).

16. 证明当且仅当 $T^2 = 1$，变换 T 的逆是自身.

17. 设变换 T 非奇异，试从定义出发，证明 $T^3T^{-5} = T^{-2}$，$(T^3)^{-5} = (T^{-5})^3 = T^{-15}$.

18. 设 T 是到直线 L 的正交射影，求证 $T^2 = T$.

19. 设 R 是绕 O 点旋转 $\frac{\pi}{3}$ 的变换. 求证 $R^6 = 1$，并证明 $R^n = 1$ 仅当 $n = 6k(k = 0, \pm 1, \pm 2\cdots)$ 时成立. R^{74}、R^{-30} 是什么？

20. 设 T 是变换，n 是使 $T^n = 1$ 成立的最小正整数，则称 T 的周期(或阶)是 n. 例如，练习 19 中的变换的周期是 6. 试求下列变换的周期：

(i) 旋转 $\frac{2}{5}\pi$.

(ii) 反射.

证明正交射影及平移不是周期变换，即不存在一个正整数 n，使 $T^n = 1$.

21. 求证周期变换是非奇异变换.

[提示：从关系式 $T^n = 1$ 中可求出 T 的逆.]

其逆成立吗？

22. 设 S、T 二个变换可以交换. 证明对任意正整数 n，$(ST)^n = S^nT^n$.

S、T 不可交换时，这个公式对吗？

6.3 平面等距变换

在 6.1 节中，我们用变换来讨论图形的对称性，这些使图形与自身重合的变换都不使图形产生变形．所以，在我们讨论一个图形的对称性时，已经遇到使任意二点距离保持不变的变换 T，亦即对任意二点 P、Q，P 与 Q 之间的距离等于它们的象 $T(P)$ 与 $T(Q)$ 之间的距离．这样的变换叫做等距变换．二个等距变换的乘积仍是等距变换．

现在我们讨论平面等距变换，从而给出它们所有可能的分类．

定理 II 平面等距变换由不在一直线上的任意三点 A、B、C 以及它们在这个等距变换下的象 A'、B'、C' 唯一确定．

证明 设 T 是等距变换，$T(A)=A'$，$T(B)=B'$，$T(C)=C'$，因为三角形 ABC 与 $A'B'C'$ 的对应边相等，所以这二个三角形全等．现在我们证明对任意一点 P，它的象（$=T(P)$）是完全确定的．因为 T 是等距变换，所以

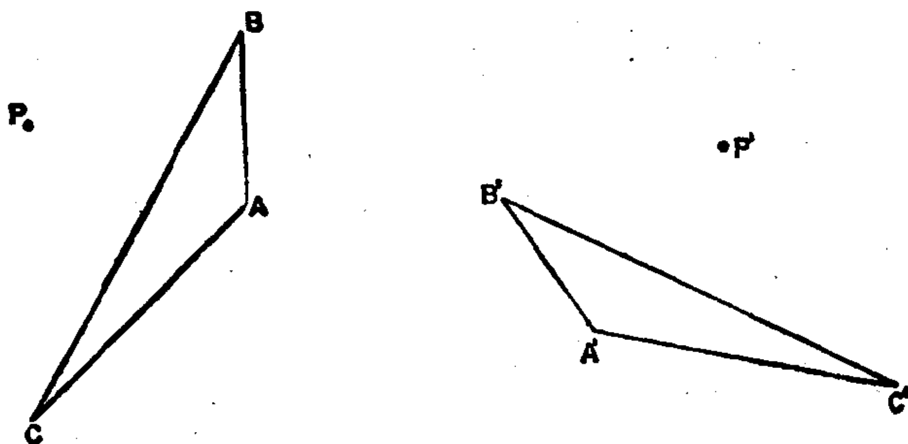


图 40 平面等距变换

$$A'P'=AP、B'P'=BP、C'P'=CP.$$

所以 P' 点在分别以 A'、B'、C' 为圆心，AP、BP、CP 为半径的三个圆上。但是，如果三个圆有二个或更多交点时，它们的圆心必然共线，所以，上面的三个圆只有唯一的交点。结论是 P' 点完全确定，这就是所要证明的。

在讨论平面等距变换时，以上定理是很有用的。这定理告诉我们，为确定一个平面等距变换，只须画一个三角形以及它在此等距变换下的象。

定理 III 只有下列四种类型的平面等距变换（图 41(a)到(d))：

(a) 旋转，

(b) 平移，

(c) 反射，

(d) 滑移反射(即反射后紧接一个平移，平移的方向沿着反射"镜面"L)。直线 L(图 41(d))称为滑移反射的轴。

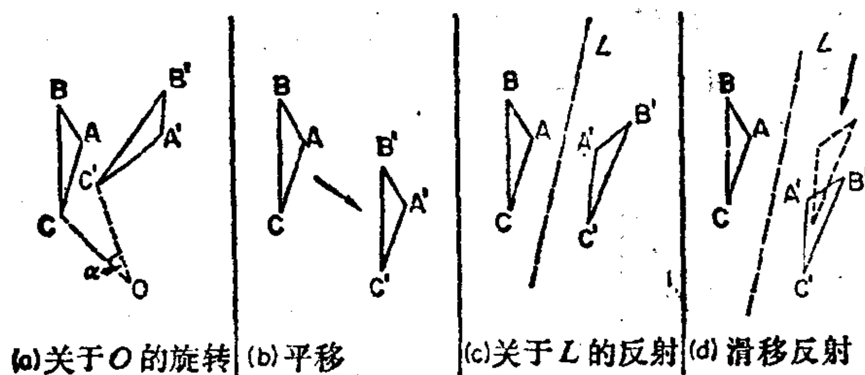


图 41 平面的四种等距变换

上面四种变换显然都是平面等距变换。我们还应证明除这四种外，再也没有其它类型的等距变换。在证明这一点前，我们先用图 41 来介绍等距变换的另一种重要的分类方法。

设 A、B、C 是不共线的三点，依次由 A 点出发作直线连接 B 点，由 B 点出发连接 C 点，由 C 点出发连接 A 点，得到一个三角形。

如果上面画三角形的过程是逆时针方向进行的，则称A、B、C三点逆时针方向；如果是顺时针方向进行的，则称顺时针方向．所以，图41(a)中的A、B、C三点逆时针方向．若三角形ABC的三个顶点按A、B、C的顺序是逆时针方向，则称这个三角形逆时针方向；同样可以定义三角形ABC顺时针方向．所以，三角形ABC与三角形ACB总是相反方向的．

如果一个等距变换把不共线的A、B、C三点映到相同方向的A'、B'、C'三点，则这个等距变换是正等距变换．不是正等距变换的等距变换叫做反等距变换．图41说明旋转与平移是正等距变换，而反射与滑移反射是反等距变换．

为了说明一个等距变换的类型，我们可以讨论ABC这个三角形移到新位置$A'B'C'$的方式．如果是正等距变换，只要将ABC三角形在平面上移动即可．如果是反等距变换，则要把ABC三角形从平面上拿起来并翻一个身才能移到新位置上．

二个反等距变换的乘积是正等距变换；二个正等距变换的乘积是正等距变换；一个正等距变换与一个反等距变换的乘积总是反等距变换，与乘积的次序无关．

现在我们能证明定理IV，它改进了定理II的结论．

定理IV 平面正等距变换由任意二个不同点A、B以及它们的象A'、B'唯一确定．对平面反等距变换也有同样的结论．

证明 设T是正等距变换，C是与A、B不共线的任意一点，C'是C在T下的象．因为三角形ABC与三角形$A'B'C'$全等而且方向相同，所以C'点唯一确定．由定理II，T唯一确定．同理可证T是反等距变换的情况．

平面等距变换的分类(定理III的证明)

情况1　T是正等距变换

设T把A映到A'，B映到B'(图42)．作AA'及BB'的垂直平分线，分别过M，N二点．设这二条垂直平分线相交于O点，则

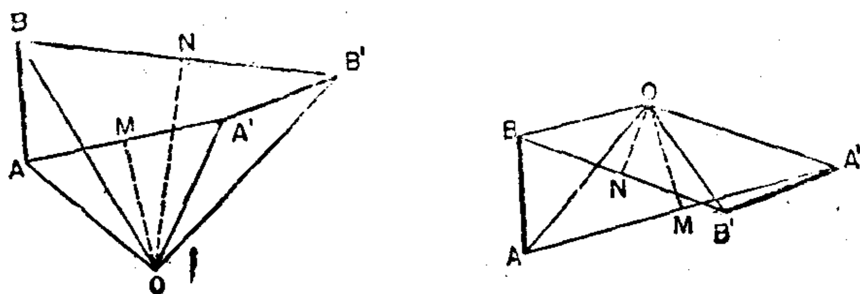


图 42　正等距变换：旋转

三角形 OAB 与三角形 $OA'B'$ 的对应边相等，所以它们是全等三角形．T 是正等距变换，因此这二个三角形方向相同．通过绕 O 点的，旋转角是角 AOA' 的旋转，可将 AB 重合到 $A'B'$ 上．由定理 IV，T 等于这个旋转．

若 AA' 及 BB' 的垂直平分线互相平行，则 $AA'B'B$ 是平行四边形或者是等腰梯形(图 43)．在第一种情况，AB 经一个平移变到 $A'B'$；在第二种情况，AB 经绕 O 点的一个旋转变到 $A'B'$．所以 T 是平移，或是旋转．

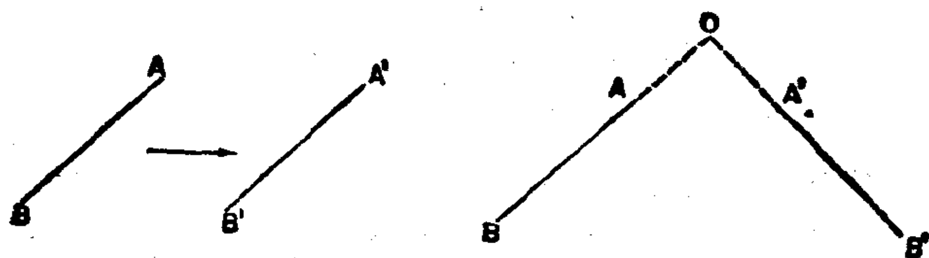


图 43

情况 2　T 是反等距变换

设 S 是任意一个反射，比如说是关于直线 L 的反射．TS 是二个反等距变换的积，所以它是正等距变换．由情况 1，TS 是平移或旋转．$S^2=1$，所以

$$T=TS^2=(TS)S.$$

上式说明T是平移与反射的乘积，或者T是旋转与反射的乘积．我们将在下面二个例子中证明，这二种乘积只可能是反射或滑移反射，这样就完成了定理 III 的证明．

例 9 设S是关于直线L的反射，A是由位置向量$\mathbf{a}$定义的平移．求证：

(i) 若$\mathbf{a}$的方向与L垂直，则AS是反射．

(ii) 若$\mathbf{a}$的方向与L不垂直，则AS是滑移反射．

证明

(i) 设P、Q是直线L上的任意不同的二点．如图 44 (a) 所示，$AS(P)=A(P)=P'$，$AS(Q)=Q'$．设L'是直线L平移$\frac{1}{2}\mathbf{a}$所得的直线，则P、Q二点关于直线L'作反射后，分别映成P'、Q'二点．由定理 IV，AS等于关于直线L'的反射．

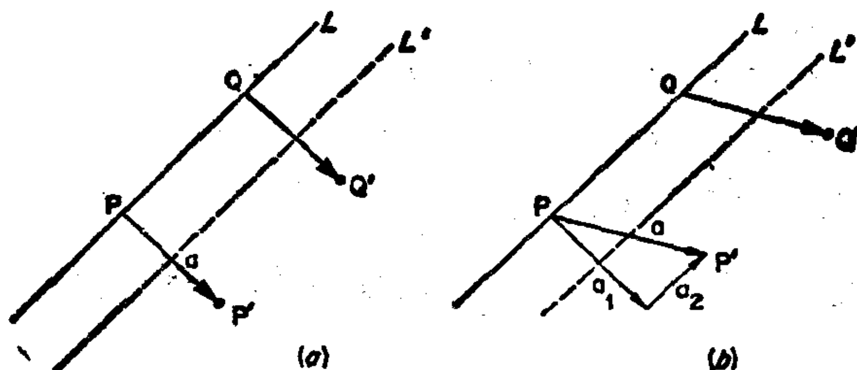


图 44 平移与反射的乘积

(ii) 将$\mathbf{a}$写成$\mathbf{a}=\mathbf{a}_1+\mathbf{a}_2$，使$\mathbf{a}_1$与$L$垂直，$\mathbf{a}_2$与$L$平行．从图 44(b) 容易看出，$AS$是关于轴$L'$的反射与$\mathbf{a}_2$所定义的平行于$L'$的平移的乘积，所以$AS$是滑移反射．

例10 设S是关于直线L的反射，R是绕O点旋转α角的旋转．求证：

(i) 若O点在直线L上，则RS是反射；

(ii) 若O点不在直线L上，则RS是滑移反射．

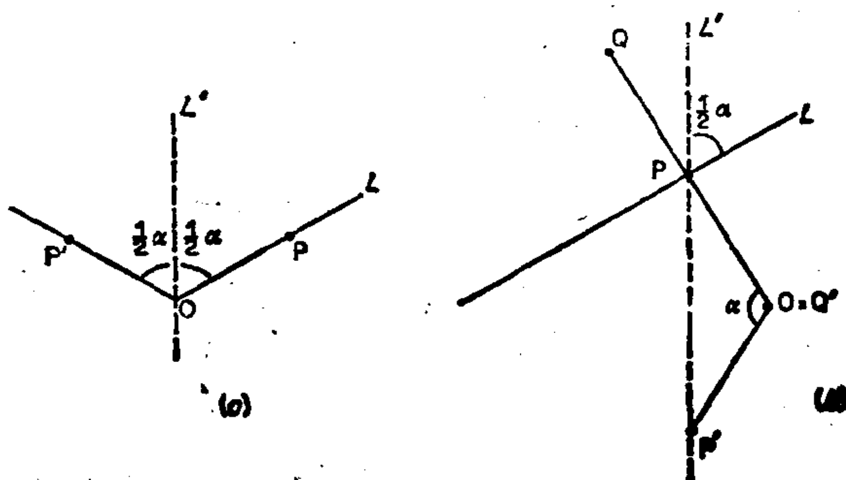


图 45 旋转与反射的乘积

(i) 设P是直线L上的另一点，则$RS(O)=O, RS(P)=P'$（图 45(a)）. 设L'与L之间的夹角是$\frac{1}{2}\alpha$，由定理 IV，RS是关于直线L'的反射.

(ii) 见图 45(b)，OPQ垂直于L，$OP=PQ$，并且，

$$RS(P)=R(P)=P', \quad RS(Q)=R(O)=O.$$

由轴L'以及位置向量$\overrightarrow{PP'}$定义的滑移反射把P、Q点映成P'、Q'点. 由定理 IV，RS等于这个滑移反射. 因为角$OPP'=\frac{1}{2}(\pi-\alpha)$，所以$L'$与$L$的夹角是$\frac{1}{2}\alpha$.

还可以根据固定点进一步对等距变换进行分类. 例如，绕O点的旋转固定O点. 下面总结了平面等距变换依此种方法进行分类的情况：

(a) 有固定点的等距变换

(i) 正等距变换：旋转，

(ii) 反等距变换：反射（反射直线上的所有点都被固定）.

(b) 没有固定点的等距变换

(i) 正等距变换：平移，

(ii) 反等距变换：滑移反射.

例 11 平移与旋转的乘积

二个正等距变换的乘积是正等距变换，所以，平移 T 与旋转 R 的乘积是旋转或平移.

设 T 由平移向量 **a** 定义，R 是绕 O 点逆时针旋转 α 角的旋转. 若 $\alpha=2k\pi$，k 是整数，则 $R=1$，所以 $TR=T$.

除此之外，TR 有一个固定点，所以它是一个旋转.

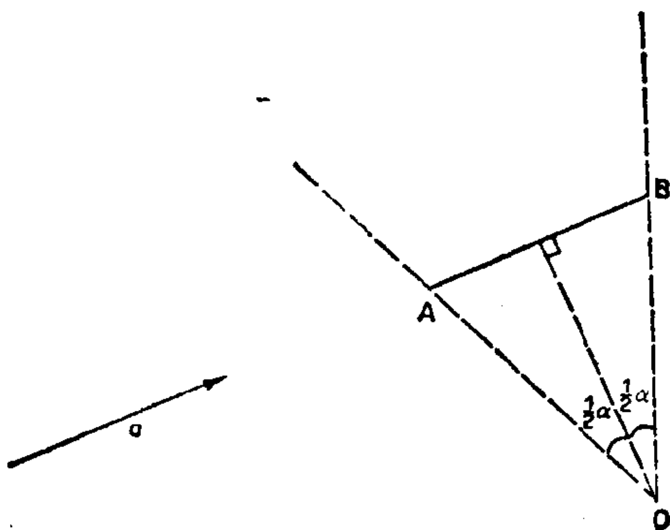


图 46 平移与旋转的乘积

图 46 的三角形 OAB 中，$\overrightarrow{AB}=\mathbf{a}$. 旋转 R 将 B 点映到 A 点，平移 T 将 A 点映到 B 点. 所以 TR（先作用 R，后作用 T）使 B 点固定. 这就是说，TR 是绕 B 点的旋转. 从 O 点在 TR 下的象可以看出 TR 是逆时针方向的旋转，旋转角为 α. 同理可证 RT（先 T，后 R）是绕 A 点的旋转，旋转角为 α. 但是，$RT\neq TR$.

等距变换一个重要性质是它们都有逆. 我们把平面等距变换及其逆列表如下：

表 8

等距变换 T	逆 T^{-1}
绕 O 点逆时针旋转 α 角	绕 O 点顺时针旋转 α 角
平移 $\mathbf{a}$	平移 $-\mathbf{a}$
关于 L 的反射	关于 L 的反射
滑移反射：关于 L 的反射以及平行于 L 的平移 $\mathbf{a}$	滑移反射：关于 L 的反射以及平移 $-\mathbf{a}$

练　习

23．S、A 如同例 9 所设．SA 是什么等距变换？

24．S、R 如同例 10 所设．SR 是什么等距变换？

25．证明例 11 的最后一个性质．

26．描述二个反射的乘积，设

(i) 二条反射直线互相平行；

(ii) 二条反射直线相交．

27．S_1、S_2 是任意二个反射．求证 $(S_1S_2)^{-1}=S_2S_1$．

28．设 G 是滑移反射．求证 G^2 是平移．

29．设 S、T 是二个可交换的等距变换，即 $ST=TS$．证明 S 与 T^{-1} 也可以交换．

30．旋转角为 π 的旋转称为半周旋转，证明二个半周旋转 R_1 与 R_2 的乘积是平移，平移的方向与这二个旋转中心的连线平行，平移的距离是二个旋转中心距离的二倍，而且

$$(R_1R_2)^{-1}=R_2R_1.$$

31．(i) 二个旋转的中心重合时，它们的乘积是什么？

(ii) 二个旋转的中心不同时，它们的乘积是什么？

在什么情况下，二个旋转可以交换？

32．设二个反射的直线互相垂直，求证它们的乘积是半周旋转．

33．试求下面二个等距变换的乘积：一个是滑移反射，一个是半周旋转，它的中心在滑移反射的轴上．这二个变换可以交换吗？

34．设一个反射的反射直线与一个滑移反射的轴垂直，求这二个变换的乘积．

6.4 三维空间等距变换

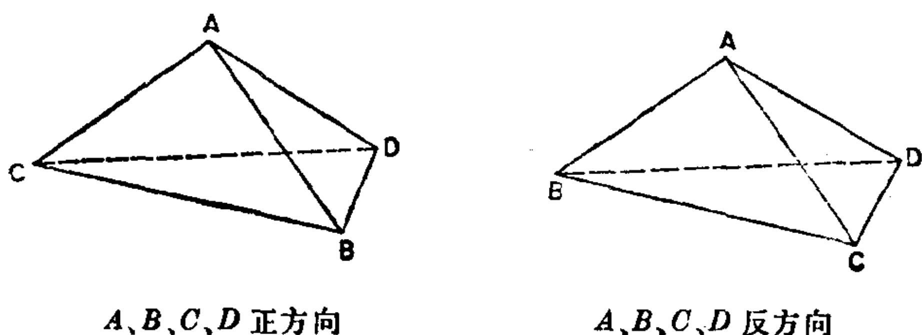
下面一个定理的证明与平面等距变换的情况类似，所以我们把它略去了.

定理 V 空间等距变换由任意不共面的四点 A、B、C、D 以及它们的象 A'、B'、C'、D' 唯一确定.

将空间等距变换分成正、反二类，也是比较方便的. 设 A、B、C、D 是空间不共面的四点，若右手螺旋绕 ABC 顺时针旋转时，此螺旋的前进方向指向（背离）D 点，则称这四点按 A、B、C、D 顺序为正（负）方向（图 47）.

如果等距变换将不共面的四点 A、B、C、D 映到同方向的四点 A'、B'、C'、D'，则这个等距变换叫做正等距变换. 不是正等距变换的等距变换称为反等距变换.

空间等距变换可由四面体 $ABCD$ 变成四面体 $A'B'C'D'$ 的变换表示出来. 仅对正等距变换，四面体 $ABCD$ 可以通过移动移到 $A'B'C'D'$ 的位置. 与平面情况一样，二个正等距变换或二个反等距变换的乘积是正等距变换，一个正等距变换与一个反等距变换的乘积是反等距变换.



A、B、C、D 正方向　　　A、B、C、D 反方向

图 47 不共面四点的方向

下面我们讨论三维空间等距变换的分类. 分类中包括伪旋转以及螺旋旋转. 我们首先讨论这二类等距变换.

在三维空间中，旋转是绕一根轴的旋转；反射是关于一个平面的反射，这个平面叫做反射平面．如果一个旋转重复几次后等于恒等变换，则它的轴称为 n 次轴．例如，旋转角为$\frac{\pi}{2}$或$\frac{3}{2}\pi$的旋转的轴是四次轴，旋转角为 π 的旋转叫做半周旋转，它的轴是二次轴．一般将旋转角是 $\frac{2}{n}\pi$ 的旋转记成 C_n（n 是整数）．显然，$(C_n)^n=1$，所以 C_n 的轴是 n 次轴．

在二维时，我们已经知道使一个点固定的旋转与反射的乘积是反射．在三维时，仅当旋转轴在反射平面中时，上述类似结果才成立．为此，可考虑绕轴 L 的旋转 R 和反射平面 π 与 L 垂直的反射 σ．乘积 $S=\sigma R$ 使 π 与 L 的交点 O 固定，但在 R 不是恒等变换时，它既不是反射，又不是旋转．这是一种新类型的变换，我们称它是关于轴 L 的伪旋转*（图 48）．

σ 与 R 是可以交换的：

$$S=\sigma R=R\sigma, \tag{12}$$

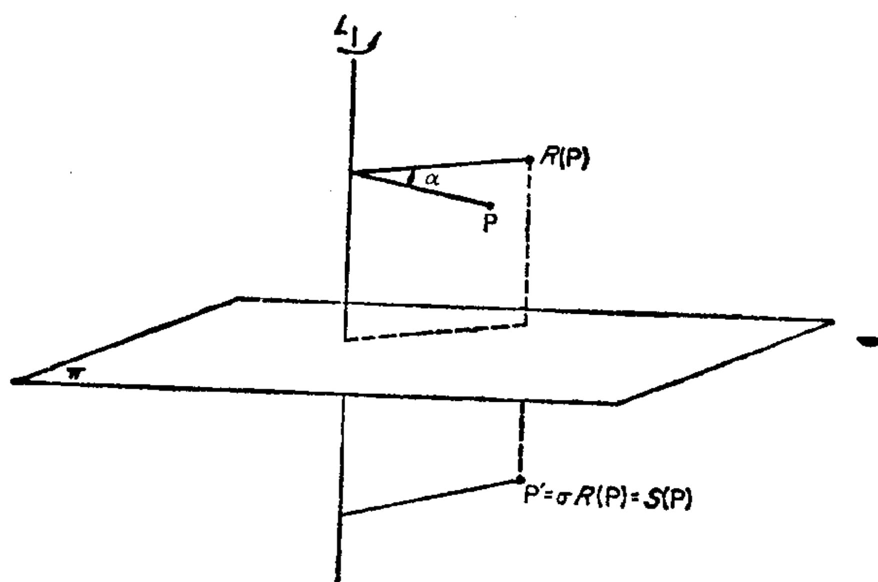


图 48　伪旋转 $S=\sigma R$

---

* 或旋转反射．

所以

$$S^{\tau}=\sigma^{\tau}R^{\tau}. \tag{13}$$

由于 $\sigma^2=1$，所以伪旋转的偶数次幂是旋转．反射使一个平面上的点固定，旋转使一条直线上的点固定，但是伪旋转仅使一个点固定：L 与 σ 的反射平面 π 的交点．

一个轴是 L 的半周旋转与一个关于平面 π 的反射的乘积，当 π 与 L 垂直时，特别重要．由这样得到的伪旋转称为关于 O 点的倒反，这里 O 点是 L 与 π 的交点．这个倒反将 P 点映到 P' 点，使 O 点是线段 PP' 的中点(图 49)．

如果取 O 点为直角坐标系的原点，则关于 O 点的倒反将点 $P(x,y,z)$ 映到点 $P'(-x,-y,-z)$．O 点叫做倒反的中心．

设 I 是倒反，则

$$I^2=1, \tag{14}$$

于是 I 是它自己的逆．

一般说来伪旋转与轴及平面有关，但倒反仅与它的中心这一个点有关．倒反的结构简单，它是几何证明中的有用工具．下面用一例说明．

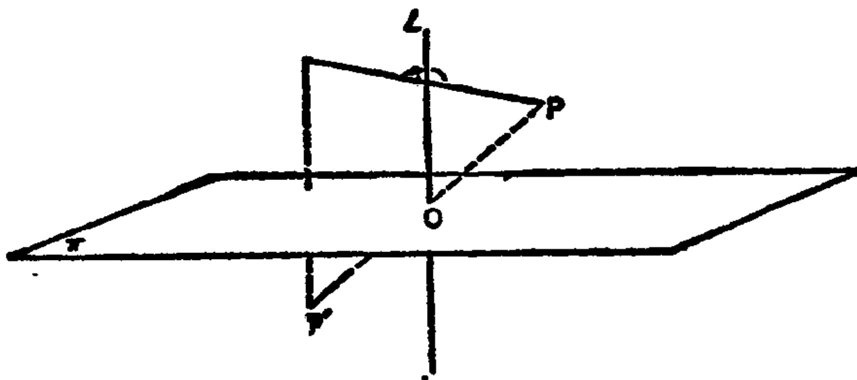


图 49　倒反

例 12　设一个正等距变换使 O 点固定，则容易证明它是以过 O 点的直线为轴的旋转（见本节的练习 37）．试证明使一点固定

的反等距变换是伪旋转.

设 S 是使 O 点固定的反等距变换， I 是以 O 点为中心的倒反. I 是反等距变换，故乘积 $IS(=R)$ 是正等距变换. R 使 O 点固定，于是 R 是以过 O 点的直线 L 为轴的旋转. 由公式(14)，

$$IR=I^2S=S.$$

将 I 写成乘积 $C_2\sigma$，C_2 是以直线 L 为轴的半周旋转，σ 是关于过 O 点与 L 垂直的平面的反射. 于是

$$S=\sigma C_2R.$$

显然，C_2R 是以 L 为轴的旋转. 所以 S 是使 O 点固定的、以 L 为轴的伪旋转.

我们要讨论的第二种等距变换是螺旋旋转. 螺旋旋转是一个旋转与一个平行于此旋转的轴的方向的平移的乘积（旋转与平移相乘的次序可以任意）. 显然，它是一个正等距变换.

例 13 证明若正等距变换没有固定点，则必是平移或螺旋旋转.

设 S 是没有固定点的正等距变换，S 不是平移. 任取一点 O，设 $S(O)=O'$. 令 T 是将 O' 点映到 O 点的平移，则 TS 是使 O 点固定的正等距变换. 于是 $TS=R$ 是一个以过 O 点的直线 L 为轴的旋转*.

把 T 写成 T_1T_2，使 T_1 是垂直于 L 的平移，T_2 是平行于 L 的平移. 这时，

$$R=T_1T_2S,$$

所以

$$T_1^{-1}R=T_2S. \tag{15}$$

T_1^{-1} 是垂直于 R 的轴的平移，因此 $T_1^{-1}R$ 是旋转. 设这个旋转是 R_1，则 R_1 的轴 L_1 与 L 平行(比较 6.3 节例 11). 由(15)，

---

* 由假设 S 不是平移，所以 R 不可能是恒等变换. 由练习 37，R 是一个旋转角不是零角的旋转，直线 L 唯一存在——译者注.

$$S=T_2^{-1}R_1,$$

这是关于 L_1 的旋转与平行于 L_1 的平移的乘积，说明 S 是螺旋旋转.

最后我们总结三维空间等距变换的分类：

(a) 有固定点的等距变换

(i) 正等距变换：旋转，

(ii) 反等距变换：伪旋转.

(b) 没有固定点的等距变换：

(i) 正等距变换：1. 平移；

2. 螺旋旋转.

(ii) 反等距变换：滑移反射.

三维滑移反射是关于平面 π 的反射与平行于 π 的平移的乘积（次序任意）. 在上面的分类中，我们将反射归入伪旋转一类，这是因为在 $S=\sigma R$ 中取 $R=1$，即可得到反射 σ.

练　　习

35. 试证明6.3节定理IV在三维空间中的推广：

任意一个正（或反）等距变换由不共线三点 A、B、C 以及它们的象 A'、B'、C' 唯一确定.

36. 求证：如果二个不同反射的反射平面互相平行，则它们的乘积是平移；否则它们的乘积是旋转，并求出所得旋转的轴和旋转角.

37. 可以如下来证明"若正等距变换 T 使 O 点固定，则 T 是旋转并且轴过 O 点"这个命题：

设 T 将三角形 OAB 映到（与其全等的）三角形 $OA'B'$.

(i) 若 $A=A'$，则我们得到一个轴为 OA，将 B 映到 B' 的旋转.（若 $B=B'$，则此旋转的旋转角是零角，即退化成一个恒等变换.）

(ii) 若 $A\neq A'$，令 M_1 是 AA' 的中点，π_1 是过直线 OM_1 并且与 OAA' 垂直的平面，证明以 π_1 为反射平面的反射 S_1 将 A 点映到 A' 点.

(iii) 设上面的反射 S_1 将 B 点映到 B^* 点. 令 π 是过直线 OA' 并且过 B^*B' 中点的平面，S 是关于 π 的反射. 则 S 使 A' 点不动并且将 B^* 点映到

B'点.

(iv) 证明旋转 SS_1 将 OAB 映到 $OA'B'$. 这个旋转的轴是 π_1 与 π 的交线，此直线过 O 点.

(v) 用练习 35 的结论完成整个证明.

38. 可用以下方法求得将 OAB 映到 $OA'B'$($A\neq A'$, $B\neq B'$)的旋转的轴：

由习题 37 (iv)，此轴在平面 π_1 上，π_1 是与 OAA' 垂直并且过 AA' 的中点 M_1的平面. 同理，此轴也在平面 π_2上，π_2是与 OBB'垂直并且过 BB'的中点 M_2的平面. 如果 π_1与 π_2 是二个不同的平面，则所求的轴就是这二个平面的交线. 对特殊情况 $\pi_1=\pi_2$，很容易找出所求的轴.

39. 如果二个旋转的轴相交，证明它们的乘积是旋转.

40. 一只右手手套在反射下的象是什么？

41. 如果三个反射的反射平面互相平行，试求它们的乘积.

42. 将一个正方体的二个相对的面都划分成相等的四小块. 其中一个面的四小块分别涂上红(R)、白(W)、蓝(B)、绿(G)四种颜色. 在正方体的其余四个面上涂白色. 如果将另一个面的四小块 1、2、3、4 分别涂上 B、G、R、W 四种颜色 (图 50)，证明此时这个涂色的正方体只有一个非平凡的对

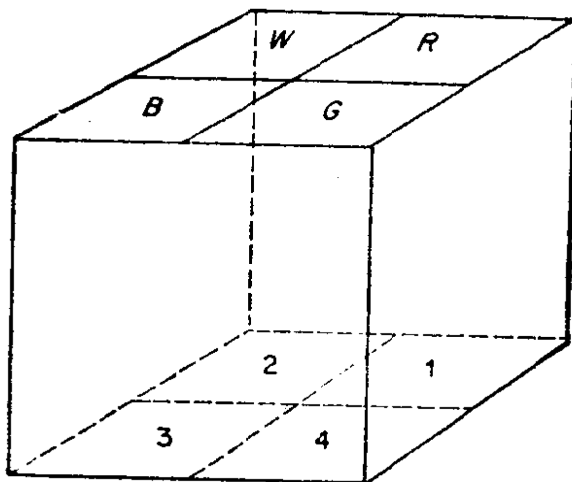


图 50 涂色正方体

称变换：以正方体中心为中心的倒反. 对以下 1、2、3、4 这四小块的三种涂色情况，试讨论此涂色正方体的对称性：

(i) G、B、W、R；

(ii) R、W、B、G；

(iii) G、R、W、B.

43. 将正方形涂上二种颜色，使涂色后的正方体有一个对称变换是四阶的伪旋转 $S_4=\sigma C_4$，而 σ 和C_4都不是对称变换.

44. 在例 13 的解答中，曾用到平移 T 可写成乘积 T_1T_2这一结论，T_1 是与 L 垂直的平移，T_2是与L 平行的平移. 试用一个草图说明.

45. 设 $S_n=\sigma C_n$ 是一个伪旋转，其中 C_n 是$\frac{2\pi}{n}$角的旋转，轴是 L；σ 是反射，它的反射平面与 L 垂直. 试求 S_n 的周期，设

(i) n 是偶数；

(ii) n 是奇数.

关于等距变换的进一步讨论可参阅：

H. S. M. Coxeter, *Introduction to Geometry*, Wiley, 1961.

第七章　群

7.1　引论

我们现在要用变换的代数给物体或图形的对称性作一个简洁的描述．本节作为引论，我们只采用二维的例子．

一个图形的对称性由其对称变换所描述，对称变换也就是这图形到自身的一个等距变换，使得变换前后的图形完全重合．

设$\mathscr{G}$是一个给定的图形(例如一个等边三角形)的所有对称变换的集合．则

(i) $\mathscr{G}$中任意两个(不一定互不相同的)变换的乘积也属于$\mathscr{G}$；

(ii) 恒等变换1属于$\mathscr{G}$；

(iii) $\mathscr{G}$中每一个T有一个逆变换(因为T是等距变换)，它仍然属于$\mathscr{G}$．

$\mathscr{G}$称为该图形的对称变换群．从下面给出的群的抽象定义来看，采用这个名称的理由是显然的．

定义 I　一个群$\mathscr{G}$是元素$x, y, z, \cdots$的集合以及一个运算$\circ$(它可以是乘法、加法或其他什么运算)，使得对$\mathscr{G}$中任意的x, y我们都能作出$x \circ y$，满足

(i) $x \circ y$是唯一定义的，并且也属于$\mathscr{G}$(封闭性公理)；

(ii) $\mathscr{G}$有一个恒等元素e，使得对$\mathscr{G}$中所有的x有

$$x \circ e = e \circ x = x;$$

(iii) $\mathscr{G}$的每一个元素x在$\mathscr{G}$中有一个逆元素，亦即，对$\mathscr{G}$的每一个x，在$\mathscr{G}$中存在一个y，使得

$$x \circ y = y \circ x = e;$$

(iv) 结合律成立：对$\mathscr{G}$中任意的x,y,z,都有

$$(x\circ y)\circ z=x\circ(y\circ z).$$

我们称$\mathscr{G}$是在运算$\circ$下的一个群.

如果$\mathscr{G}$只有有限个 元素，则称为有限群. 群$\mathscr{G}$中 元素的个数称为这个群的阶,记作$|\mathscr{G}|$.

我们看到，一个图形的对称群确实是在取变换乘积这个结合运算下的一个群. 还有群的其他许多例子. 让我们暂时离开主题来举几个例子吧.

例 1

所有非零有理数$\frac{m}{n}$(m 和 n 是$\neq 0$ 的整数)的集合在乘法下成为一个群. 验证群的四条公理如下：

(i) 封闭性：两个非零有理数的积也是一个非零 有理数.

(ii) 恒等元素就是数 1.

(iii) $\frac{m}{n}$的逆元素是$\frac{n}{m}$.

(iv) 乘法是结合的.

例 2

所有有理数的集合在加法下成为一个群. 验证如下.

(i) 封闭性：两个有理数的和也是有理数.

(ii) (在加法下的)恒等元素就是数 0,因为对所 有 的 有理数 r,

$$r+0=0+r=r.$$

(iii) r 的(加法的)逆元素是$-r$.

(iv) 加法是结合的：对任意有理数 x,y,z,总有

$$(x+y)+z=x+(y+z).$$

例 3

所有整数 $0,\pm 1,\pm 2,\cdots$的集合在加法下成为一个群. 这是

所有有理数在加法下所成的群(例 2)的一个子群.

一般地说，群$\mathscr{G}$的一个子群$\mathscr{H}$是$\mathscr{G}$的元素的一个集合，它在与群$\mathscr{G}$相同的运算$\circ$下自己成为一个群. 我们记作$\mathscr{H}\subseteq\mathscr{G}$. 如果$\mathscr{H}$是$\mathscr{G}$的一个子群而并不含有$\mathscr{G}$的所有元素，则$\mathscr{H}$是$\mathscr{G}$的一个真子群(记作$\mathscr{H}\subset\mathscr{G}$).

例 4

设$\mathscr{V}$是一个向量空间. $\mathscr{V}$在向量加法下是一个群! 这里恒等元素是零向量，$\mathbf{v}$ 的逆元素是$-\mathbf{v}$.

例 5

复数 $1,-1,i,-i$ 在乘法下成为一个有限群.

一个有限群的元素合成的方式能够列一张群表而描述出来. 这是一张正方形的表格，它的行与列分别用群的不同元素标明. 元素 $x\circ y$ 出现在以 x 开头的行和以 y 开头的列的交叉处，如下面的群表(i)所示.

	e	a	b	$\cdots$	y	$\cdots$
e	e	a	b	$\cdots$	y	$\cdots$
a	a					
b	b					
$\vdots$	$\vdots$				$\vdots$	
x	x				$x\circ y$	
$\vdots$	$\vdots$					

(i)

	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

(ii)

习惯上第一行与第一列标以恒等元素 e. 此时，这一行和这一列只是简单地重复所标明的群的元素. 对于例 5，完整的群表由(ii)给出.

从一个群表容易验证封闭性公理——表中必须只含有群的元

素．下例中封闭性公理就不成立了．元素 a,b 与 c 在下表所定义的运算下不成为一个群：

	a	b	c
a	b	a	d
b	a	c	b
c	c	b	a

一个群表中每一行和每一列含有群的每个元素一次并且只有一次．

证明 假设由 x 开头的一行含有群的元素 g 两次．则存在群的互不相同的元素 y 与 z，使得

$$g=x\circ y=x\circ z.$$

设 x^{-1} 表示 x 的逆元素．我们有

$$x^{-1}\circ(x\circ y)=x^{-1}\circ(x\circ z).$$

于是，由结合律，

$$(x^{-1}\circ x)\circ y=(x^{-1}\circ x)\circ z.$$

因而 $e\circ y=e\circ z$，即 $y=z$．与假设矛盾，于是对于行的证明就完成了．对于列，可类似地证明．

群的元素的逆元素从群表中容易读出．为了求 x 的逆元素，只要找到以 y 开头的列，使它包含以 x 开头的行的恒等元素 e．此时 $x\circ y=e$，即 y 是 x 的逆元素．

上面例 1 到例 5 都是交换群．亦即，交换律

$$x\circ y=y\circ x（对所有\ x,y）$$

成立的群．下面是一个非交换群的例子．

例 6 等边三角形的对称变换群

我们在 6.1 节表 5 中列举过等边三角形的对称变换．这六个变换组成这三角形的对称变换群．群的元素是

1：恒等元素；

R_1：关于中心逆时针方向旋转$\frac{2}{3}\pi$；

R_2：关于中心逆时针方向旋转$\frac{4}{3}\pi$；

S_i：关于 L_i 的反射($i=1,2,3$)。

它的群表如下：

	1	R_1	R_2	S_1	S_2	S_3
1	**1**	R_1	R_2	S_1	S_2	S_3
R_1	R_1	R_2	**1**	S_3	S_1	S_2
R_2	R_2	**1**	R_1	S_2	S_3	S_1
S_1	S_1	S_2	S_3	**1**	R_1	R_2
S_2	S_2	S_3	S_1	R_2	**1**	R_1
S_3	S_3	S_1	S_2	R_1	R_2	**1**

我们来验证表中的某些细节．我们知道，

$$R_2=R_1^2, \qquad R_1^3=\mathbf{1}.$$

用这二个等式就可以验证表的左上角部分．例如，

$$R_2R_1=(R_1^2)R_2=R_1^3=\mathbf{1}$$

以及

$$R_2^2=(R_1^2)^2=R_1^4=R_1.$$

我们再考虑一个旋转和一个反射的乘积．用 6.3 节例10的方法，可求得

$$R_1S_1=S_3, \qquad S_1R_1=S_2.$$

等边三角形的对称变换群不是交换群．显然，当且仅当一个群的群表关于主对角线对称时，这个群才是一个交换群．

元素 $\mathbf{1}, R_1$ 与 R_2 组成上面这个群的一个子群．元素 $\mathbf{1}$ 与 S_1 也是如此．你能找出别的子群吗？

由于等边三角形的对称变换群的六个元素都可以表成两个元素 R_1 与 S_1以及它们的逆元素的乘积：

$$1=R_1R_1^{-1}, \qquad R_2=R_1^2, \qquad S_2=S_1R_1, \qquad S_3=S_1R_1^2,$$

我们说这个群是由 R_1 与 S_1 生成的. 一般地, 说一个(有限或无限)群 $\mathscr{G}$ 是由 $g_1, g_2, \cdots, g_r, \cdots$ 生成的,即 $\mathscr{G}$ 的任意元素都可由有限个 g_i 及其逆元素(允许重复)合成. 群的元素 $g_1, \cdots, g_r, \cdots$ 称为 $\mathscr{G}$ 的生成元. 把这个定义与一个向量空间的生成元的定义作比较,有一个重要的不同点. 在一个向量空间中,不但生成元的和*属于这个空间,而且它们的纯量倍也属于这个空间. 由单个元素生成的群称为循环群.

例 7 证明正方形的对称变换群由两个元素生成.

提示:证明它的每一个对称变换都可表成有限个旋转与反射的乘积,其中的旋转角是 $\frac{\pi}{2}$,反射的直线是正方形一边的中垂线. 这个群有八个元素,其中四个是反射.

例 8 正多边形的对称变换群:平面二面体群(例 6 与例 7 的推广).

一个正 n 边形($n=8$ 的情况如图 51 所示)有旋转对称:

$$C_n: \text{关于中心 } O \text{ 逆时针旋转 } \frac{2}{n}\pi.$$

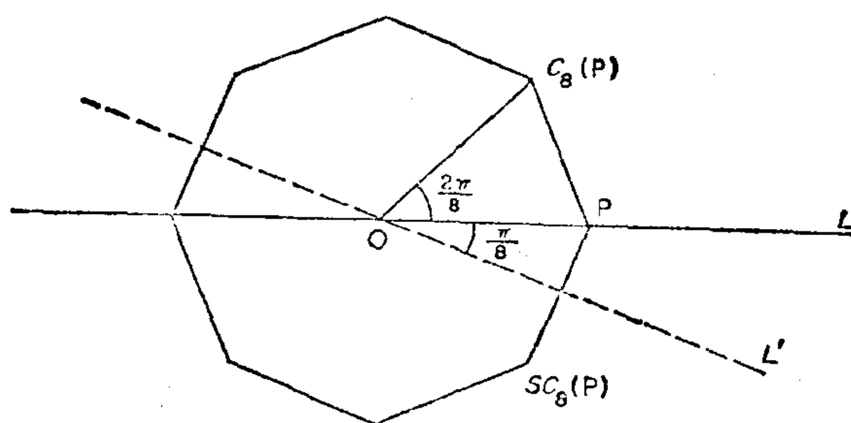


图 51 正八边形的对称交换群

* 这个运算是向量的加法.

C_n 生成了所有旋转对称的子群 $\mathscr{C}_n$. 这个群是 n 阶的；其不同的元素是 $C_n^r(r=0,1,2,\cdots,n-1)$. 旋转对称群的记号 $\mathscr{C}_n$ 是通用的. 字母$\mathscr{C}$表示$\mathscr{C}_n$是一个循环群. 这个多边形还有反射对称

S：关于连接 O 点与多边形一个顶点的直线 L 的反射.

整个对称群是由 C_n 与 S 生成的. 它是 $2n$ 阶的群，它的不同元素为

$$C_n^r \text{ 与 } SC_n^r \quad (r=0,1,\cdots,n-1) \tag{1}$$

元素 SC_n^r 是反射. 特别，SC_n 是关于一条与 L 交成$\frac{\pi}{n}$角的直线 L'的反射.

生成元 C_n 与 S 满足关系式

$$S^2=1, \quad (C_n)^n=1, \quad C_nS=SC_n^{1-}. \tag{2}$$

这三个关系式足以判别这些生成元及其逆元素的任意乘积究竟是群里的哪个元素了；这样的乘积总可以化简成（1)中所列的形式之一. 由于这个原因，关系式（2)称为这个群的定义关系式.

正 n 边形的对称变换群经常被称为二面体群，记成$\mathscr{D}_n$. 这个术语不太适当，因为它还被用于横截面是正 n 边形的正棱柱体

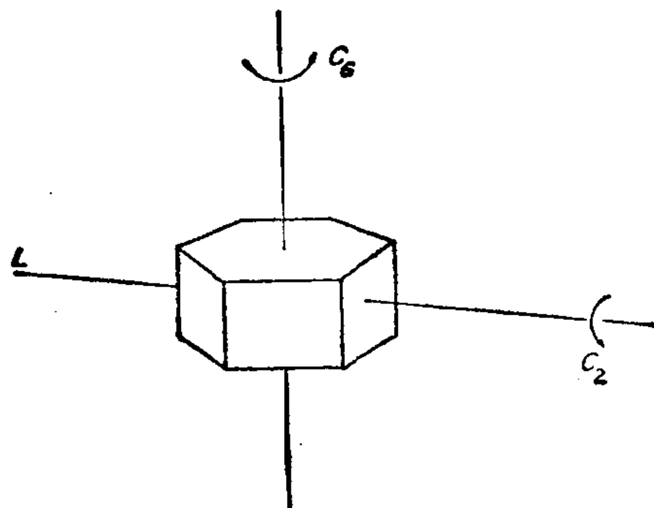


图 52 正六棱柱的旋转对称

($n=6$ 的情况如图 52 所示）的旋转对称的群．这个群是由关于一条 n 次旋转轴的旋转 C_n 以及关于一条 2 次旋转轴的半周旋转 C_2 所生成．如果我们考虑一个无限薄的棱柱，并且把棱柱上的点当作平面上的点，则上述平面二面体群和三维二面体群之间的类似之处就成为十分明显的了．关于 L 的半周旋转在这些点上的作用就是平面上关于 L 的反射的作用．

以后我们还将看到，平面二面体群 $\mathscr{D}_n$ 与群 $\mathscr{C}_{nv}$(8.5 节的例 3 与例4)之间所具有的相似之处，比它与其三维同名者之间的相似性更强．

练　　习

1．以下各例中哪些是群？哪些是循环群？

(i) 整数 0，±1，±2，…，在减法下；

(ii) 偶整数 0，±2，±4，…，在加法下；

(iii) 偶整数在乘法下；

(iv) 所有满足 $z^8=1$ 的复数的集合，在乘法下．(这些数即所谓八次单位根．)

2．证明在由 S 与 C_n 生成的二面体群 $\mathscr{D}_n$ 中，

(i) $C_n^{-1}=(C_n)^{n-1}$，

(ii) $(SC_n)^{-1}=SC_n$．

试求 C_n^r 与 SC_n^r 的逆元素的类似表达式．

3．在平面二面体群 $\mathscr{D}_4$ (正方形的对称变换群）中，利用定义关系式 (2) 化简下列乘积．

(i) C_4SC_4，

(ii) $C_4^2SC_4^2$，

(iii) $SC_4^{-1}SC_4^2S$．

在下列的练习题中，群是用乘法写出的，亦即，群的两个元素的合成 $a\cdot b$ 简记成 ab．

4．证明在一个群中消去律成立：如果 a、b、g 都是群的元素，且 $ag=bg$，则 $a=b$．

5．如果 a、b、c、d 与 g 都是群 $\mathscr{G}$ 的元素，使得 $agb=cgd$，由此能够推出 $ab=cd$ 吗？

［提示：参考练习 3 (i) 与(ii)．］

证明如果 $\mathscr{G}$ 是交换群，则可以消去 g.

6．设 $\mathscr{G}$ 是一个图形的对称变换群，下面各例中哪些构成 $\mathscr{G}$ 的子群？

(i) 所有旋转对称的集合（注意：恒等变换看成平凡的旋转）.

(ii) 所有平移对称连同恒等变换.

(iii) 所有反射对称连同恒等交换.

(iv) 所有正等距变换.

(v) 所有反等距变换.

7．群的元素 g 的阶是使 $g^n=e$（恒等元素）的最小正整数 n. 证明一个有限群的所有元素都是有限阶的.

［提示：考虑幂 g，g^2，g^3，$\cdots$. 在一个有限群中它们不可能都是互异的.］

8．设 g 是 12 阶的群元素$\left(\text{例如，}\frac{\pi}{6}\text{的旋转}\right)$.下列元素各是几阶的？

$$\text{(i) } g^2;\ \text{(ii) } g^3;\ \text{(iii) } g^4;\ \text{(iv) } g^5;\ \text{(v) } g^8.$$

试推广所得的结论.

9．$\mathscr{H}$ 和 $\mathscr{K}$ 都是 $\mathscr{G}$ 的子群. $\mathscr{H}$ 和 $\mathscr{K}$ 的交（记成 $\mathscr{H}\cap\mathscr{K}$）定义为 $\mathscr{H}$ 与 $\mathscr{K}$ 的所有公共元素的集合. 证明 $\mathscr{H}\cap\mathscr{K}$ 是 $\mathscr{G}$ 的子群.

10．作出正方形的对称变换群 $\mathscr{D}_4$ 的乘法表. 试把这个群的子群分类.

11．设 $\mathscr{G}$ 是一个群，它的恒等元素是 e. "如果 $a^2=e$ 并且 $b^2=e$，则 $(ab)^2=e$"这一断言在下述情况下正确与否？

(i) 在交换群 $\mathscr{G}$ 中；

(ii) 在所有的群 $\mathscr{G}$ 中.

［提示：考虑二面体群的元素.］

12．找出一个非循环的交换群.

7.2 群的同构

现在我们再来看上节曾介绍过的二个群：一个是 $\mathscr{C}_4$，它是绕 O 点旋转 $\frac{\pi}{2}$ 的整数倍角的旋转的群；另一个是 $\mathscr{G}$，它是由复数 1、

-1、i、$-i$ 组成的乘法群.

$\mathscr{G}$ 由 i 生成（也可由 $-i(=i^3)$ 生成）. 群 $\mathscr{G}$ 的不同元素是：

$$i、i^2、i^3、i^4=1.$$

群 $\mathscr{C}_4$ 的不同元素是

$$C_4、C_4^2、C_4^3、C_4^4=1.$$

将 $\mathscr{G}$ 与 $\mathscr{C}_4$ 比较，发现除了记号上的不同外，它们显然是相同的. 我们称它们是同构的.

一般来说，若二个群 $\mathscr{G}$ 与 $\mathscr{G}'$ 除了记号上可能有不同外，它们是相同的，也就是在它们的元素之间可找到一个一一对应 $g\longleftrightarrow g'$，使得如果

$$g\longleftrightarrow g' \quad 与 \quad h\longleftrightarrow h',$$

就有

$$g\circ h\longleftrightarrow g'\circ' h',$$

那末称 $\mathscr{G}$ 及 $\mathscr{G}'$ 同构，记成 $\mathscr{G}\cong\mathscr{G}'$. 保持结构的一一对应称为群的同构. 上式中我们假设 $\mathscr{G}$ 的运算是 $\circ$，$\mathscr{G}'$ 的运算是 $\circ'$. 这二个运算可以是不同的. 例如在上例中，一个是复数乘法，一个是变换的乘法.

上面讨论的两个群都叫做 4 阶循环群的实现或模型. 当然，数字 ±1、$\pm i$ 与旋转 $\frac{\pi}{2}$ 的整数倍是数学中的完全不同的量. 但是这二个群有相同的群的结构.

为了比较不同群的群的结构，采用只反映群的结构性质而不涉及其它性质的中性符号来抽象地写出群所包含的元素，是一个方便方法. 我们可以把 4 阶循环群抽象地记成

$$gp\{c\,|\,c^4=e\}^*,$$

读成"由元素 c 生成的群，c 满足关系式 $c^4=e$."（e 是恒等

* 也常写成 $\mathscr{G}=\{c\,|\,c^4=e\}$——译者注.

元）.

因此，在同构的意义下只有一个 n 阶循环群，可记成*

$$gp\{c \mid c^n = e\}.$$

下面是 n 阶循环群的两个（同构的）模型：

(i) 旋转 $\frac{2\pi}{n}$ 的整数倍角的旋转组成的群 $\mathscr{C}_n$.

(ii) 复数

$e^{2\pi ri/n} = \cos(2\pi r/n) + i\sin(2\pi r/n)$ $(r=0, 1, \cdots, n-1)$ 的乘法群，

因为这个群的元素恰是方程

$$z^n = 1$$

的所有根，所以这个群也叫做 n 次单位根群.

复数 ± 1、$\pm i$ 的群与旋转所组成的群 $\mathscr{C}_4$ 之间的同构使我们联想起众所周知的复数的几何解释.

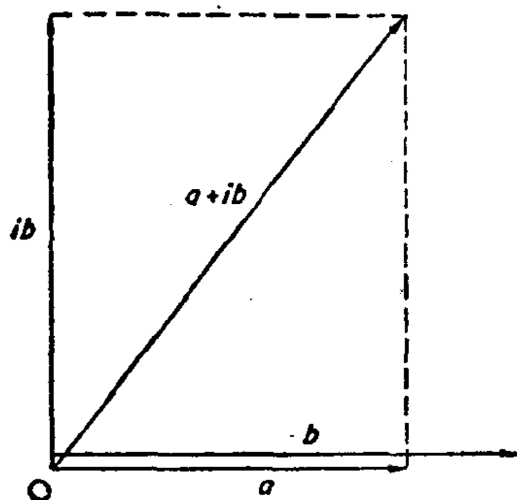


图 53 阿冈德图

取一个平面直角坐标系，用水平轴上离原点距离为 b 的点来代表实数 b，也可等价地用这点的位置向量来代表这个数（图 53）. 现在我们把 ib 解释成"将 b 点的位置向量逆时针旋转 $\frac{\pi}{2}$ 角". 复数 $a+bi$ 是相加二个位置向量而得到的数. 这就是大家已熟悉的阿冈德图. 上面说明了阿冈德图的几何意义. 显然复数 $i(a+ib)$ 由阿冈德图中的 $a+bi$ 的位置向量旋转 $\frac{\pi}{2}$ 而得到. 一般地，用复数 $e^{2\pi i/n}$ 相乘相当于作用一个逆时针旋转 $\frac{2\pi}{n}$ 的旋转.

---

* 在这里我们采用乘法记号. 若用加法记号，可写成 $gp\{c \mid nc=0\}$. 一般地讲，"∘"这个记号很不确切.

练　习

13. 证明交换群不可能和非交换群同构.

14. 证明二个不同阶的群不能同构.

15. 证明由绕一个固定轴旋转$\frac{2\pi}{6}$的整数倍角的旋转所成的群$\mathscr{C}_6$与二面体群$\mathscr{D}_3$不同构. 此例说明二个阶数相同的群不一定同构.

16. 设$\mathscr{G}$与$\mathscr{G}'$是二个同构的群，它的同构是保持结构的一一对应$g\leftrightarrow g'$. 求证：

(i) $e\leftrightarrow e'$,这里e与e'分别是$\mathscr{G}$与$\mathscr{G}'$的恒等元;

(ii) 若$g\leftrightarrow g'$,则$g^{-1}\leftrightarrow (g')^{-1}$;

(iii) 若$g\leftrightarrow g'$,则$g^r\leftrightarrow (g')^r$对一切整数$r$成立,并且元素$g$与$g'$有相同的阶.

17. 验证下面变换的集合都是正方形的对称变换群$\mathscr{D}_4$的子群:

(i) C_4、C_4^2、C_4^3、$\mathbf{1}$;

(ii) S、C_4^2、SC_4^2、$\mathbf{1}$;

(iii) SC_4、SC_4^3、C_4^2、$\mathbf{1}$.

证明其中有二个是互相同构的群,而另一个与它们不同构.

7.3　置换

本节我们要讨论置换. 我们将看到,置换在对称变换群的理论中很重要,在行列式的理论中也很重要.

考虑由n个不同元素组成的一个集合,为了方便,将这些元素记成1、2、$\cdots n$. 如果π是使这个集合中的每个元素对应于这个集合中的一个元素的一一变换,则称π是这个集合的一个置换*. 所以,π把1映到a_1,2映到a_2,$\cdots$,n映到a_n,使$a_1,a_2,\cdots a_n$是符号$1,2,\cdots n$的一个排列. 如果

$$\pi(i)=a_i\ \ (i=1,2,\cdots n) \tag{3}$$

* π是置换的通用符号,这种用法不会与弧度π发生混淆.

则把这个置换记成

$$\pi=\begin{pmatrix}1 & 2\cdots & n\\ a_1 & a_2\cdots & a_n\end{pmatrix}. \tag{4}$$

下面一行中的每个元素是它上面的元素在 π 作用下的象.

例如,当 $n=5$ 时,置换

$$\pi=\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 3 & 5 & 1 & 4 & 2\end{pmatrix}, \tag{5}$$

把 1 变到 3, 2 变到 5 等等. 这时可写

$$\pi(1)=3,\ \pi(2)=5,\ \pi(3)=1,\ \pi(4)=4,\ \pi(5)=2.$$

我们也可把这个置换等价地写成

$$\begin{pmatrix}2 & 1 & 4 & 3 & 5\\ 5 & 3 & 4 & 1 & 2\end{pmatrix} \text{ 或 } \begin{pmatrix}5 & 4 & 1 & 3 & 2\\ 2 & 4 & 3 & 1 & 5\end{pmatrix} \text{ 等等;}$$

在上面的每一种写法中,下面一行的符号都是它顶上的符号在 π 下的象. 每个置换可以唯一地写成标准形,即上面一行写成递升序列 1,2,3···. 而有时将置换写成非标准形比较方便.

n 个符号的置换称为 n 阶置换.可以证明恰有 $n!=n(n-1)\cdot(n-2)\cdots2\cdot1$ 个不同的 n 阶置换. 请注意

$$\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 4 & 3\end{pmatrix}$$

不是置换. 这个变换把符号 2 与 3 都变成 4,所以它不是一一的变换.

置换的乘积

两个 n 阶的置换 π_1 与 π_2 的乘积定义成

$$\pi_1\pi_2(i)=\pi_1(\pi_2(i)) \quad (i=1,2,\cdots,n). \tag{6}$$

乘积 $\pi_1\pi_2$ 的作用是先作用 π_2 后作用 π_1*. 例如,

---

* 在有些书上作用的先后恰好相反。

$$\pi_1=\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix},\quad \pi_2=\begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix},$$

则

$$\pi_1\pi_2=\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix}=\begin{pmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{pmatrix}.$$

在求乘积 $\pi_1\pi_2$ 时，要记住从右往左读出置换. 所以，1 在 $\pi_1\pi_2$ 下的象是 2，它可用下法求得：

$$1\xrightarrow{\pi_2}1\xrightarrow{\pi_1}2.$$

同样可以求得

$$2\xrightarrow{\pi_2}3\xrightarrow{\pi_1}1,\quad 3\xrightarrow{\pi_2}2\xrightarrow{\pi_1}3.$$

可以看出

$$\pi_2\pi_1=\begin{pmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{pmatrix}\neq\pi_1\pi_2.$$

这说明置换的乘法不满足交换律. 但是，置换的乘法确实满足结合律.

证明 只要把 6.2 节定理 1 中作用在平面的点上的变换换成作用在符号上的变换即可.

恒等置换

恒等置换是置换

$$\begin{pmatrix}1 & 2\cdots i\cdots n\\ 1 & 2\cdots i\cdots n\end{pmatrix}.$$

我们将把恒等置换记成 $\mathbf{1}$. 对任意的置换 π，显然有 $\mathbf{1}\,\pi=\pi\,\mathbf{1}=\pi$.

逆置换

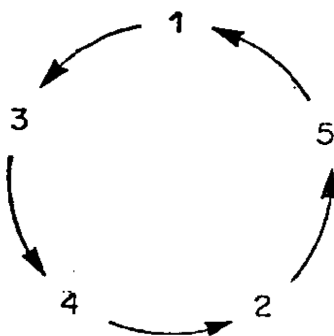
置换有一个重要的性质：每个置换都有逆置换.

置换 $\pi=\begin{pmatrix}1 & 2\cdots & n\\ a_1 & a_2\cdots & a_n\end{pmatrix}$的逆置换是 $\pi^{-1}=\begin{pmatrix}a_1 & a_2 & \cdots & a_n\\ 1 & 2 & \cdots & n\end{pmatrix}$.

这是因为显然有 $\pi^{-1}\pi=\pi\pi^{-1}=1$.

循环

如果一个置换作用在一个集合上的方式能用类似下面的图表达出来.



则这个置换称为循环. 图中表示的置换可以简记成(13425),这样写时,每个符号后面紧接的是它的象,而最后一个符号的象是第一个符号. 这个置换还可以写成下面四种形式:(34251),(42513),(25134)或(51342). 含有 n 个符号的循环叫做长度是 n 的循环或 n-循环.

任意一个置换可方便地写成循环的记法. 例如,置换

$$\pi=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 6 & 7 & 5 & 4 & 1 & 3 & 2\end{pmatrix}.$$

π 将 1 变到 6, 6 变到 3, 3变到5, 5 变到 1. 所以循环(1635)是 π 的一个组成部分. π 的另一个组成部分是 (27),剩下的符号 4 在 π 下不变. 显然 π 可以写成循环的乘积 $\pi=(1635)(27)(4)$在(1635),(27),(4) 这三个循环中,每个字母最多只出现一次,这样的循环称为不相交的循环. 不难证明任意一个置换可以写成不相交的循环的乘积.

在把一个置换写成循环的乘积时，往往省略长度是1的循环．所有没有出现的符号应理解成在此置换下固定不变．例如，上面的置换 π 可写成 $\pi=(1635)(27)$．请注意 $(1635)(27)=(27)(1635)$．任意两个不相交的循环在相乘时可以交换．但是二个含有相同符号的循环，一般是不能交换的．例如，$(123)(145)=(14523)$，但 $(145)(123)=(12345)$．在求上面的乘积时，要记住我们曾约定从右向左乘．

练　习

18．把下列置换用循环的记号表出：

(i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 1 & 5 & 7 & 2 & 6 & 4 \end{pmatrix}$；

(ii) $\begin{pmatrix} 2 & 4 & 3 & 1 & 6 & 7 & 5 \\ 4 & 1 & 3 & 2 & 7 & 5 & 6 \end{pmatrix}$．

19．把下列循环写成不相交循环的乘积：

(i) $(1\ 2\ 3)(1\ 4\ 7\ 6)(2\ 3)$；(ii) $(1\ 5)(1\ 4)(1\ 3)(1\ 2)$；

(iii) $(1\,a)(1\,b)(1\,a)$．

20．证明循环 $(b_1b_2\cdots b_n)$ 的逆循环是 $(b_n\cdots b_2b_1)$．

21．设 $\pi_1,\pi_2,\cdots,\pi_r$ 是置换．求证

$$(\pi_1\pi_2\cdots\pi_r)^{-1}=\pi_r^{-1}\cdots\pi_2^{-1}\pi_1^{-1}.$$

举例说明 $(\pi_1\pi_2)^{-1}=\pi_1^{-1}\pi_2^{-1}$ 这个式子是错的．

22．求练习18的置换的逆置换，并求以下置换的逆置换：

(i) $(1\ 2\ 3)(1\ 4\ 5)$；(ii) $(2\ 3)$；(iii) $(1\ 4)(2\ 3)$．

23．设 $\pi=(1\ 2\ 3\ 4\ 5\ 6)$，求 π^2,π^3 等等．π 的周期等于什么？更一般地，试找出一个 n-循环的周期．

24．设 π 如同练习23，$\pi'=(7\ 8\ 9\ 10)$，$\pi^*=(7\ 8\ 9)$．试求 $\pi\pi'$ 及 $\pi\pi^*$ 的周期．一般地，如果 π 与 π' 是不相交循环，而且周期分别是 m 及 n，试求 $\pi\pi'$ 的周期．

25．设 $\pi=(1\ 2)$，$\pi'=(1\ 3)$．试求 $\pi\pi'$ 的周期并与练习24比较．

26．证明共有 $4!(=24)$ 个不同的4阶置换．

27．验证置换 $1,(1\ 2)(3\ 4),(1\ 3)(2\ 4),(1\ 4)(2\ 3)$ 组成一个

四阶交换群．这个交换群是循环群吗？

7.4 置换在对称变换群中的应用

一个图形的对称变换群常常可以方便地表示成由置换组成的群．考虑正方形的对称变换群．我们用数字 1、2、3、4 来代表正方形的四个角(图 54)．如果对称变换将 i 角变到 a_i角，那末我们用置换

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ a_1 & a_2 & a_3 & a_4 \end{pmatrix}$$

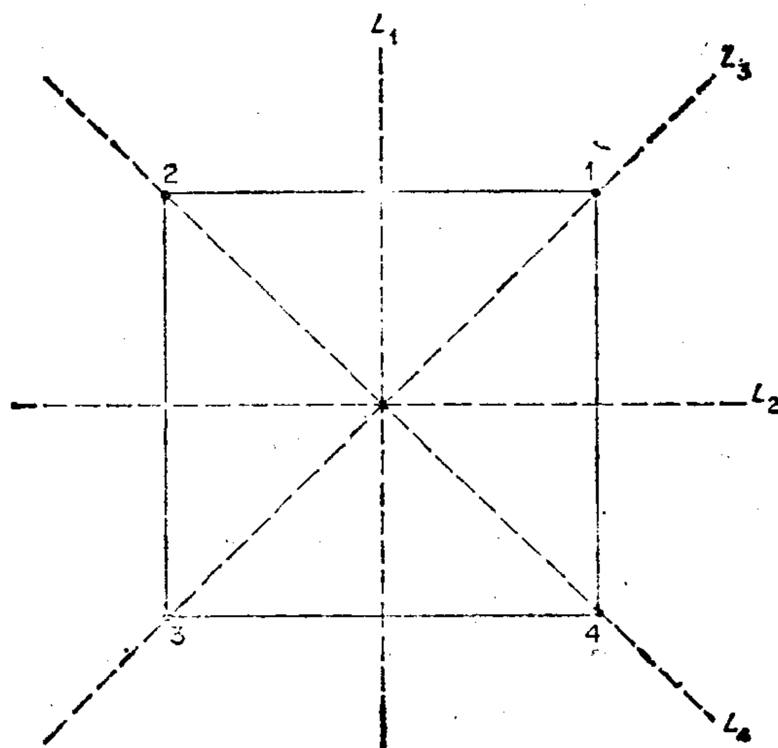


图 54 正方形的对称变换群

来表示这个对称变换． 标号 1、2、3、4 应看成是一个固定的参照系．这时，描述正方形的对称性的每个等距变换都可以表示成为一个 4 阶置换，并且二个对称变换的乘积被表示成对应的置换的乘积．

表 9 列举了正方形的对称．

表 9

对　称　变　换	置换表示(写成循环的乘积)
1. C_4，绕中心逆时针旋转$\frac{\pi}{2}$	(1 2 3 4)
2. C_4^2	(1 3)(2 4)
3. C_4^3	(1 4 3 2)
4. C_4^4=恒等变换	1
5. S_1，关于 L_1 的反射	(1 2)(3 4)
6. S_2，关于 L_2 的反射	(1 4)(2 3)
7. S_3，关于 L_3 的反射	(2 4)
8. S_4，关于 L_4 的反射	(1 3)

正方形的对称变换群由 C_4 与 S_1 生成，如果用置换的语言讲，它由(1 2 3 4)与(1 2)(3 4)生成．容易验证对称变换的乘积对应于它们所对应的置换的乘积．例如，$C_4S_1=S_4$ 与相对应的关系式(1 2 3 4)(1 2)(3 4)=(1 3)是一致的．

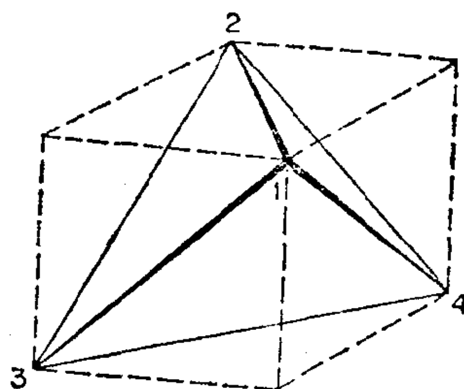


图 55　正四面体内接于正方体

四面体群

四面体群是正四面体的旋转对称变换构成的群．一个正四面体可以内接于一个正方体(图 55)．

把这个正四面体的四个顶点标上 1、2、3、4 这四个符号．对这个固定的参照系，四面体的对称变换可以表示成 4 阶置换．共有 $4!(=24)$个4阶置换，其中一半表示旋转，这一半组成的群$\mathscr{T}$就是所谓的四面体群．这是 12 阶群，它由以下两个旋转生成：其一是关于过顶点 1 的一根三次轴、旋转角是$\frac{2}{3}\pi$ 的旋转，它对应于

置换(2 3 4)；其二是半周旋转，它的轴是过正方体中心的铅垂方向的直线，它对应于置换(1 2)(3 4).

另外的 12 个置换也表示此正四面体的对称变换，但不是旋转对称. 例如，置换(1 2)表示一个反射，它是关于过 3、4 二点的铅垂方向平面的反射. 又如置换(1 4 2 3)表示伪旋转 $\sigma_h C_4^z$，它是关于过正方体中心 O 的铅垂方向轴的、$\frac{\pi}{2}$的旋转，再接一个关于过 O 点的水平平面的反射的合成. 如果 P 点在位置 1 上，P' 点在位置 4 上，则显然这个旋转将 P 点映到 P' 点. 同样，它将位置 4 上的点 Q 映到位置 2 上的点 Q'，等等.

正四面体的完全对称变换群有 24 个元素. 通常记成 $\mathscr{T}_d$. 如上所见，$\mathscr{T}_d$ 与由四个符号的所有置换构成的群同构. 12 个旋转对称成为 $\mathscr{T}_d$ 的子群 $\mathscr{T}$. 非旋转对称全体不成群. 它们既不满足封闭性公理，其中又没有恒等变换. 四面体分子甲烷(CH_4)的对称变换群是 $\mathscr{T}_d$.

练　　习

28. 把下面图形的对称变换群表示成置换群：

(i) 等边三角形；

(ii) 正六边形(用 6 阶置换).

29. 如图 56，在正方体的顶点处标上 1、2、···8. 证明旋转对称 C_4^x（绕 Ox 轴按图所指方向旋转$\frac{2}{4}\pi$)可表示成置换(1 2 3 4)(5 6 7 8). 找出旋转对称 C_4^z 对应的置换. 证明乘积 $C_4^x C_4^z$ 是绕一个三次轴的旋转.

30. 求正方体的旋转对称变换群的阶.

31. 写出四面体群的 12 个元素.

32. 用图 55 中的符号，试确定下列置换所表示的对称变换的类型（旋转、反射或伪旋转)：

(i) (1 4 2)；

(ii) (1 3)；

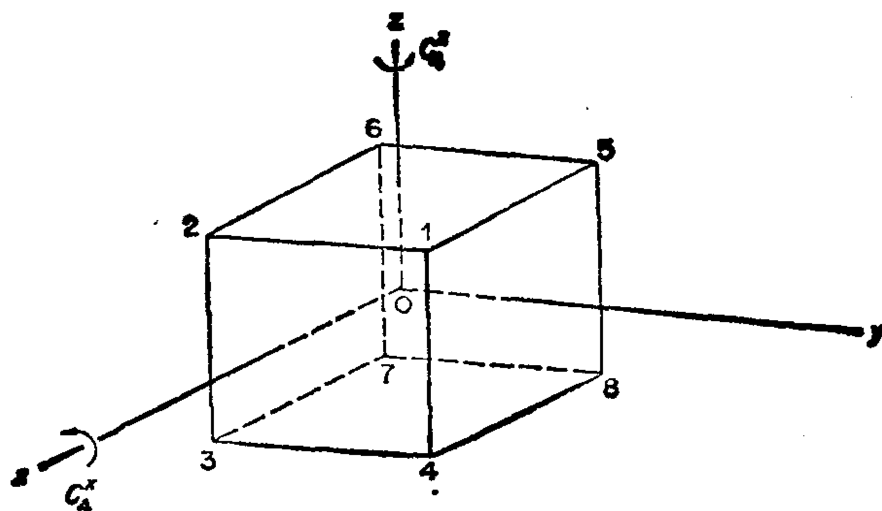


图 56 正方体的旋转对称

(iii) (2 4);

(iv) (1 3)(2 4);

(v) (1 3 2 4);

(vi) (1 2 3 4).

33. 多于一个变量的多项式的对称性可以由使它不变的变量的置换来刻划. 例如,将 x_1及 x_2对换而使 x_3及 x_4 固定这个置换使多项式 $x_1x_3+x_1x^4+x_2x_3+x_2x_4$ ($=(x_1+x_2)(x_3+x_4)$) 不变. 我们把这个置换写成(1 2), 并且称它是这个多项式的对称变换. 在下列置换中,哪些是这个多项式的对称变换?

(i) (1 2)(3 4); (ii) (1 2 3 4); (iii) (1 3 2 4);

(iv) (3 4); (v) (1 3)(2 4); (vi) (1 2 3).

证明表示对称变换的所有置换集合构成一个群. 它称为该多项式的对称变换群.

34. 求以下多项式的对称变换群:

(i) $x_1^2-2x_2^2+x_3^2$,

(ii) $x_1x_2+x_2x_3+x_3x_4+x_4x_1$.

35. 试找出二个以上的 x_1、x_2及 x_3的多项式, 使它们的对称变换群由所有的 3 阶置换构成. 这样的多项式称为三个变量的对称多项式. 它包含了所有可能的对称. 它们的对称变换群叫做三阶对称群.

36. 证明正方形的对称变换群与多项式$(x_1+x_3)(x_2+x_4)$的对称变换群

同构.

7.5 偶置换与奇置换

本节我们进一步讨论置换的性质，它在行列式理论中有重要的应用.

一个 2-循环 $\tau=(i\,j)$是一类特别简单的置换. 这种置换叫做对换. 它把 i 变到 j,把 j 变到 i,而使没有出现的符号保持不变. 请注意,$\tau^2=1,\tau=\tau^{-1}$.

现在我们要证明任意的 n-循环可以写成 $(n-1)$ 个对换的乘积. 我们可以直接计算下面(7)式的右边，来证明(7)式成立:

$$(a_1a_2\cdots a_n)=(a_1a_n)\cdots(a_1a_3)(a_1a_2). \tag{7}$$

(7)式右边的乘积次序很重要,我们是从右向左相乘,所以 $a_1\to a_2$, $a_2\to a_1\to a_3$, $a_3\to a_1\to a_4$等等. (7) 式右边的对换有一个公共的字母 a_1,所以相乘次序是不能交换的.

可以有多种方法把一个 n-循环写成对换的乘积. 例如,

$(1\ 3\ 4\ 2\ 5)=(1\ 5)(1\ 2)(1\ 4)(1\ 3)$, 但也可以写成

$(1\ 3\ 4\ 2\ 5)=(3\ 4\ 2\ 5\ 1)=(3\ 1)(3\ 5)(3\ 2)(3\ 4)$.

并且各种写法中所含有的置换因子的个数也可以不等. 比如, $(1\ 3\ 4\ 2\ 5)=(1\ 5)(1\ 2)(1\ 3)(1\ 3)(1\ 4)(1\ 3)$是六个对换的乘积. 但是，我们立即要证明一个置换不可能既能写成偶数个对换的乘积又能写成奇数个对换的乘积.

因为每个置换都能写成循环的乘积，所以我们得到下面的重要结论.

定理 I 每个置换都能写成对换的乘积.

例 9 把置换

$$\pi=\begin{pmatrix}1&2&3&4&5&6&7\\7&3&6&2&5&4&1\end{pmatrix}$$

写成对换的乘积.

先将 π 写成循环的乘积：$\pi=(1\ 7)(2\ 3\ 6\ 4)$. 用(7)式，得 $\pi=(1\ 7)(2\ 4)(2\ 6)(2\ 3)$.

例 10 π 如同例 9 所设，把 π^{-1} 写成对换的乘积.

由 6.2 节公式(6.8)，$\pi^{-1}=(2\ 3)^{-1}(2\ 6)^{-1}(2\ 4)^{-1}(1\ 7)^{-1}$. 由于对换的逆是自身，所以 $\pi^{-1}=(2\ 3)(2\ 6)(2\ 4)(1\ 7)$.

一般来说，如果一个置换写成对换乘积是

$$\pi=\tau_1\tau_2\cdots\tau_m, \tag{8}$$

则它的逆置换可以写成

$$\pi^{-1}=\tau_m\cdots\tau_2\tau_1. \tag{9}$$

定义 II 根据一个置换写成偶数个或奇数个对换的乘积，称这个置换是偶置换或奇置换.

可以设想一个置换既是偶置换又是奇置换. 例如，例 9 中的偶置换为什么不能写成五个对换的乘积呢？下面的定理排除了这种可能性.

定理 II 任何一个置换不可能既是偶置换又是奇置换.

证明 设作用在 n 个符号 1、2、$\cdots n$ 上的置换 π 可以同时写成偶数个对换的乘积及奇数个对换的乘积. 由(8)，(9)二式，π^{-1} 也有同样的性质，所以恒等变换$(=\pi\pi^{-1})$可以写成奇数个对换的乘积.

考察对换 $\tau=(i\,j)$ 在下面多项式上的作用

$$\left.\begin{aligned} p(x_1,\cdots,x_n)=(x_1-x_2)(x_1-x_3)(x_1-x_4)\cdots(x_1-x_n)\\ \times(x_2-x_3)(x_2-x_4)\cdots(x_2-x_n)\\ \star(x_3-x_4)\cdots(x_3-x_n)\\ \cdots\\ \times(x_{n-1}-x_n)\end{aligned}\right\} \tag{10}$$

可以证明 τ 不是这个多项式的对称变换(见 7.4 节练习 33)，它将 $p(x_1\cdots x_n)$ 变成 $-p(x_1\cdots x_n)$. 所以 k 个对换的乘积将 $p(x_1\cdots x_n)$ 变成 $(-1)^k p(x_1\cdots x_n)$. 但恒等变换使这个多项式不变，所

以恒等变换不能写成奇数个对换的乘积．与上段矛盾，定理证毕．

多项式(10)称为 $x_1,\cdots,x_n$ 的交错多项式．

例 11 验证对换(2 3)使 x_1、x_2、x_3 的交错多项式乘上一个负号．

解 把 $p(x_1,x_2,x_3)=(x_1-x_2)(x_1-x_3)(x_2-x_3)$中的 x_2 及 x_3对换，那末可以得到$(x_1-x_3)(x_1-x_2)(x_3-x_2)$，它等于$-p(x_1, x_2, x_3)$．

n 是奇数时，n-循环是偶置换；n是偶数时，n-循环是奇置换．置换的乘法显然满足下面的法则：

$$偶\times偶=奇\times奇=偶,$$

$$偶\times奇=奇\times偶=奇.$$

于是我们由置换(132579)(468)是奇置换与偶置换的乘积，立即知道这是一个奇置换．

恒等置换是偶置换，它可以写成二个对换的乘积(12)(12)．

一个置换群中的所有偶置换的集合是一个子群．例如，正四面体的完全对称变换群$\mathscr{T}_d$作为一个 4 阶置换的群，有一个偶置换全体组成的子群——四面体群(见 7.4 节)

练　　习

37．下面二个置换哪个是奇置换，哪个是偶置换？

$$\text{(i)}\ \begin{pmatrix}1&2&3&4&5&6&7\\3&1&7&6&4&5&2\end{pmatrix},$$

$$\text{(ii)}\ \begin{pmatrix}1&2&3&4&5&6&7\\3&7&1&6&4&5&2\end{pmatrix}.$$

38．证明任意的 n-阶置换可以写成(12)，(13)，$\cdots$(1 n)中的一些对换的乘积．这就证明了所有 n 阶置换的群由(12)，(13)，$\cdots$(1 n)生成．

39．证明在置换群里，所有偶置换的集合是一个子群．所有奇置换的集合是子群吗？试用等边三角形以及正方形的对称变换群来说明．

40．验证任意对换把交错多项式 $p(x_1, x_2,x_3,x_4)$ 变成$-p(x_1,x_2,x_3,$

x_4)，求证这个交错多项式的对称变换群由所有 4 阶偶置换组成．这个群称为 4 阶交错群．

7.6 群的直积

设 $\mathscr{G}$ 有二个子群 $\mathscr{H}$ 与 $\mathscr{K}$，使

(i) $\mathscr{G}$ 由 $\mathscr{H}$ 与 $\mathscr{K}$ 的元素生成，

(ii) $\mathscr{H}$ 的每个元素与 $\mathscr{K}$ 的每个元素可以交换，即

$$hk=kh$$

对 $\mathscr{H}$ 中的任意元素 h 及 $\mathscr{K}$ 中的任意元素 k 成立．

(iii) $\mathscr{H}$ 与 $\mathscr{K}$ 仅有一个公共元素——恒等元素 e，即

$$\mathscr{H}\cap\mathscr{K}=gp\{e\},$$

则称 $\mathscr{G}$ 是 $\mathscr{H}$ 与 $\mathscr{K}$ 的直积*，记成 $\mathscr{G}=\mathscr{H}\times\mathscr{K}$．直积是讨论群的性质以及讨论群的分类的一个重要工具．我们在下一章还要讨论群的直积．

下面我们要举二个不是直积的例子，以消除可能产生的误解．

例 12

设 $\mathscr{G}=gp\{c\,|\,c^{12}=e\}$．$\mathscr{G}$ 是 12 阶循环群（你可以把这个群看成是由旋转 $\frac{\pi}{6}$ 的倍数角的旋转组成的群）．令 $\mathscr{H}$ 是由 c^2 生成的子群，$\mathscr{K}$ 是由 c^3 生成的子群．因为 $c=(c^3)(c^2)^{-1}$，所以 $c^r=(c^3)^r(c^2)^{-r}$，这里 c^{3r} 在 $\mathscr{K}$ 中，c^{-2r} 在 $\mathscr{H}$ 中，可见 $\mathscr{G}$ 由 $\mathscr{H}$ 与 $\mathscr{K}$ 的元素生成．$\mathscr{H}$ 中的每个元素与 $\mathscr{K}$ 中的每个元素可交换．但 $\mathscr{G}$ 不是 $\mathscr{H}$ 与 $\mathscr{K}$ 的直积，因为 $\mathscr{H}$ 与 $\mathscr{K}$ 的公共元素除了 e 外，还有 c^6 这个元素，条件(iii)不满足．

如果 $\mathscr{L}$ 取成由 c^4 生成的群，则 $\mathscr{G}=\mathscr{L}\times\mathscr{K}$．

---

* 在用加法记号时，我们要求对所有的 h,k 有 $h+k=k+h$．此时，$\mathscr{G}$ 叫做 $\mathscr{H}$ 与 $\mathscr{K}$ 的直和，记成 $\mathscr{G}=\mathscr{H}+\mathscr{K}$——原注．

现在通常写成 $\mathscr{G}=\mathscr{H}\oplus\mathscr{K}$——译者注．

例 13

设 $\mathscr{G}$ 是等边三角形的对称变换群；$\mathscr{H}$ 是旋转对称组成的子群，它由 C_3 生成；$\mathscr{K}$ 是由反射 S 生成的子群.

$\mathscr{G}$ 由 $\mathscr{H}$ 与 $\mathscr{K}$ 的元素生成，而且 $\mathscr{H}\cap\mathscr{K}$ 中只有一个恒等元素. 但是，$\mathscr{G}$ 不是 $\mathscr{H}$ 与 $\mathscr{K}$ 的直积，因为 $\mathscr{H}$ 中的元素 C_3 与 $\mathscr{K}$ 中的元素 S 不可交换.

例 14

直积 $\mathscr{H}\times\mathscr{K}$ 中，$\mathscr{H}$ 或 $\mathscr{K}$ 可以是非交换群. 直积的条件(ii)与 $\mathscr{H}$（或 $\mathscr{K}$）中的二个元素不交换是不矛盾的.

定理 III 设 $\mathscr{G}=\mathscr{H}\times\mathscr{K}$. 则

(a) $\mathscr{G}$ 的每个元素 g 可以唯一写成 $g=hk$，使 h 属于 $\mathscr{H}$，k 属于 $\mathscr{K}$；

(b) $|\mathscr{G}|=|\mathscr{H}||\mathscr{K}|$.

证明 (a)首先我们要证明 $\mathscr{G}$ 的每个元素 g 可以写成 hk 的形式，h 在 $\mathscr{H}$ 中，k 在 $\mathscr{K}$ 中. 由直积的条件(i)，g 可以写成有限个 $\mathscr{H}$ 中的元素 $h_1,\cdots h_r$ 及 $\mathscr{K}$ 中的元素 $k_1,\cdots k_s$ 的乘积. 由条件(ii)，在 k_j 后面的任意元素 h_i 可以移到 k_j 的前面，而不改变乘积的结果. 例如，若 $g=h_1k_1h_2k_2h_3$，则也有 $g=(h_1h_2h_3)(k_1k_2)$. 经过数次交换位置，g 显然可以写成 hk 的形式.

为证明唯一性，设 $g=hk$ 并且 $g=h_1k_1$，这里 h,h_1 是 $\mathscr{H}$ 中的元素，k,k_1 是 $\mathscr{K}$ 中的元素. 则

$$h_1^{-1}(hk)k^{-1}=h_1^{-1}(h_1k_1)k^{-1}$$

所以

$$h_1^{-1}h=k_1k^{-1}.$$

上式左边是 $\mathscr{H}$ 中的元素，右边是 $\mathscr{K}$ 中的元素. 所以两边都属于 $\mathscr{H}\cap\mathscr{K}$. 由直积的条件(iii)，

$$h_1^{-1}h=e=k_1k^{-1}.$$

于是 $h=h_1$，$k=k_1$. 这就证明了 g 的表达式的唯一性.

证明 (b)在乘积hk中，h有$|\mathscr{H}|$种可能，k有$|\mathscr{K}|$种可能，所以共有$|\mathscr{H}||\mathscr{K}|$个这样的乘积. 由证明(a)，这些乘积取遍$\mathscr{G}$的所有元素，而且它们都不相同. 这就证明了关于$\mathscr{G}$的阶的公式.

例 15

在例 12 中，$\mathscr{K}$是 4 阶的，$\mathscr{L}$是 3 阶的，$\mathscr{G}=\mathscr{L}\times\mathscr{K}$是 $3\times 4=12$ 阶的群.

练习

41. 证明循环群 $\mathscr{G}=gp\{c|c^{15}=e\}$可以写成一个 5 阶循环子群与一个 3 阶循环子群的直积.

42. 一个 12 阶循环群能写成一个 2 阶子群与一个 6 阶子群的直积吗?

43. 验证四面体群$\mathscr{T}$由代表旋转的两个置换(234)与(12)(34)生成（参照图55所标的数字）. 令$\mathscr{H}$与$\mathscr{K}$分别是由(234)及(12)(34)生成的$\mathscr{T}$的子群. 求证$\mathscr{T}$的元素不一定都能写成hk的形式，使h在$\mathscr{H}$中，k在$\mathscr{K}$中.

第八章　对称变换群

8.1　平面对称变换群

我们现在要讨论平面上重复图案的对称变换群，即平面对称变换群．这种群也叫做二维空间群．它们在晶体学中有重要意义．晶体由原子的有规则的三维排列组成．所以，用任意平面去截晶体或作晶体的投影都得到平面上重复图案，这些图案可以近似看成是无限延伸的．

我们先看几个平面对称变换群的例子．假定我们讨论的图案复盖整个平面，我们这里只讨论这种图案的对称变换群．

我们讨论的图案有一个共同点：它们都含有一个图案单元，这个图案单元以某种规律在平面上重复出现．为了说明图案单元如何重复，我们要考察这个图案的对称变换群的平移与滑移反射所生成的子群．我们首先讨论平移及滑移反射．旋转对称与反射对称可以很简单地用旋转的中心及反射的轴来表示．与对称变换群的滑移反射轴一样，这些中心与轴也随着图案单元的重复而有规律地重复．

平移与滑移反射

图 57 中所有的图案的平移对称所成的（交换）子群由二个平移生成．这二个平移由图 57 中的二个位置向量 $\mathbf{a}_i$，$\mathbf{a}_i'$ 决定，分别记为 T_i,T_i'．

图 57（1）的图案单元显然可以取成一个平行四边形．在图案的对称变换群中的平移作用下，这个图案单元布满整个平面．图 57(2)是它的一个特例，此时，图案单元是一个矩形．

图 57(3 a)及(3 b) 的图案单元是一个六边形．请注意 T_3 及

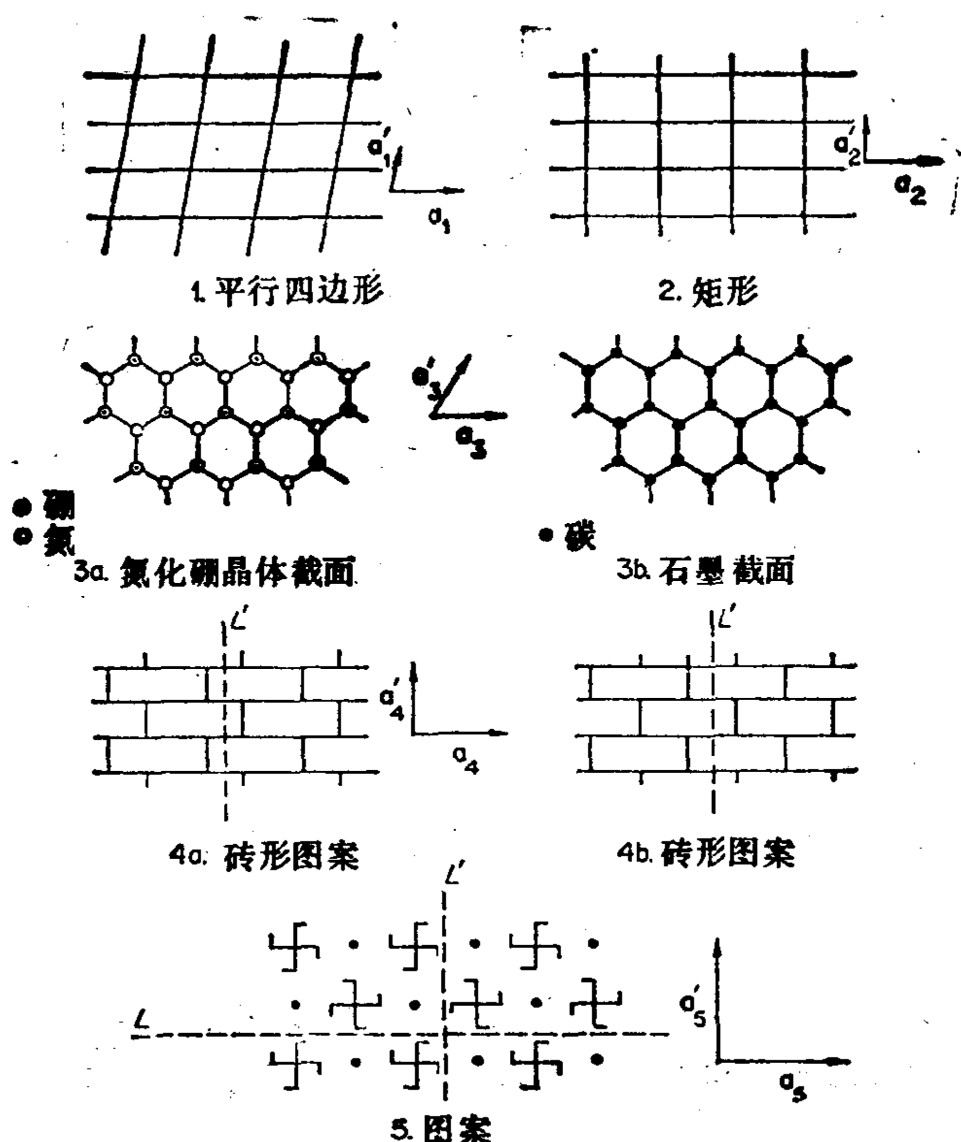


图 57 图案的一些例子

T_3' 把一个图案单元移到与它相邻的二个单元上. 可以分别通过平移 $T_3^{-1}T_3'$, T_3^{-1}, $(T_3')^{-1}$, $T_3(T_3')^{-1}$ 使图案单元移到其余四个相邻单元上. 这四个平移分别由位置向量 $\mathbf{a}_3'-\mathbf{a}_3$, $-\mathbf{a}_3$, $-\mathbf{a}_3'$, $\mathbf{a}_3-\mathbf{a}_3'$ 决定.

对图 57 (4 a)及(4 b), 我们自然地把某一块砖作为图案单元. 仅仅通过平移的作用,单独一块砖不能布满整个(无限)图案, 每二层中有一层要漏掉. 如果一个图案的对称变换群中含有基本滑移反射, 即一个不能写成这个图案的平移对称变换与反射对称

变换的乘积的滑移反射，那末它都具有这种特性．这里 $T_4'=(G_4')^2$，其中 G_4' 是基本滑移反射对称变换，它由一个关于 L' 的反射以及一个 $\frac{1}{2}\mathbf{a}_4'$ 的平移组成．单独一块砖在 T_4 及 G_4' 的乘积的作用下布满整个图案．请注意，元素 $T_4^i G_4'$ 是关于平行于 L' 的轴的一些滑移反射．每块砖上有二根这样的轴通过．

图 57(5) 的对称变换群含有水平方向与铅垂方向的基本滑移反射．（它也包含与水平成 $\pm\frac{\pi}{4}$ 方向的基本滑移反射，这些基本滑移反射与旋转 $\frac{\pi}{2}$ 的对称有关，见练习 7）．$T_5=(G_5)^2$，$T_5'=(G_5')^2$，这里 G_5 及 G_5' 分别是关于轴 L 及轴 L' 的基本滑移反射．在 G_5 与 G_5' 的乘积的作用下，图案单元卍·布满整个无限的图案．

旋转与反射

接着我们来讨论上面各个图案的旋转对称与反射对称．图58上画出了一些旋转对称与反射对称．前面我们已经讲过，$\frac{2\pi}{n}$ 的旋转对称的中心是 n 次中心．在图 58 中，小透镜◯代表一个 2 次中心（即半周旋转的中心），三角形△、正方形□、六边形⬡分别代表三次、四次、六次中心．本例没有其他类型的 n 次中心．只要画出一个图案单元的旋转对称的中心以及反射对称的轴（用虚线）就够了．在整个图案中，这些对称变换有规律地重复出现．

在图案(1) 中，平行四边形状的图案单元的中心，每个顶点及每边中点都是 2 次中心(即半周旋转中心)．

在图案(3 b)中，六角形状的图案单元的中心是 6 次中心，各个顶点是 3 次中心．这个图案还有一些 2 次中心．但在图案(3 a)中，因为有二种类型的原子，所以比图案(3 b)少一些对称．并且，把硼原子变成氮原子的旋转与反射不是这个图案的对称变换，因此六边形的中心仅是一个 3 次中心．剩下几个图案的旋转对称

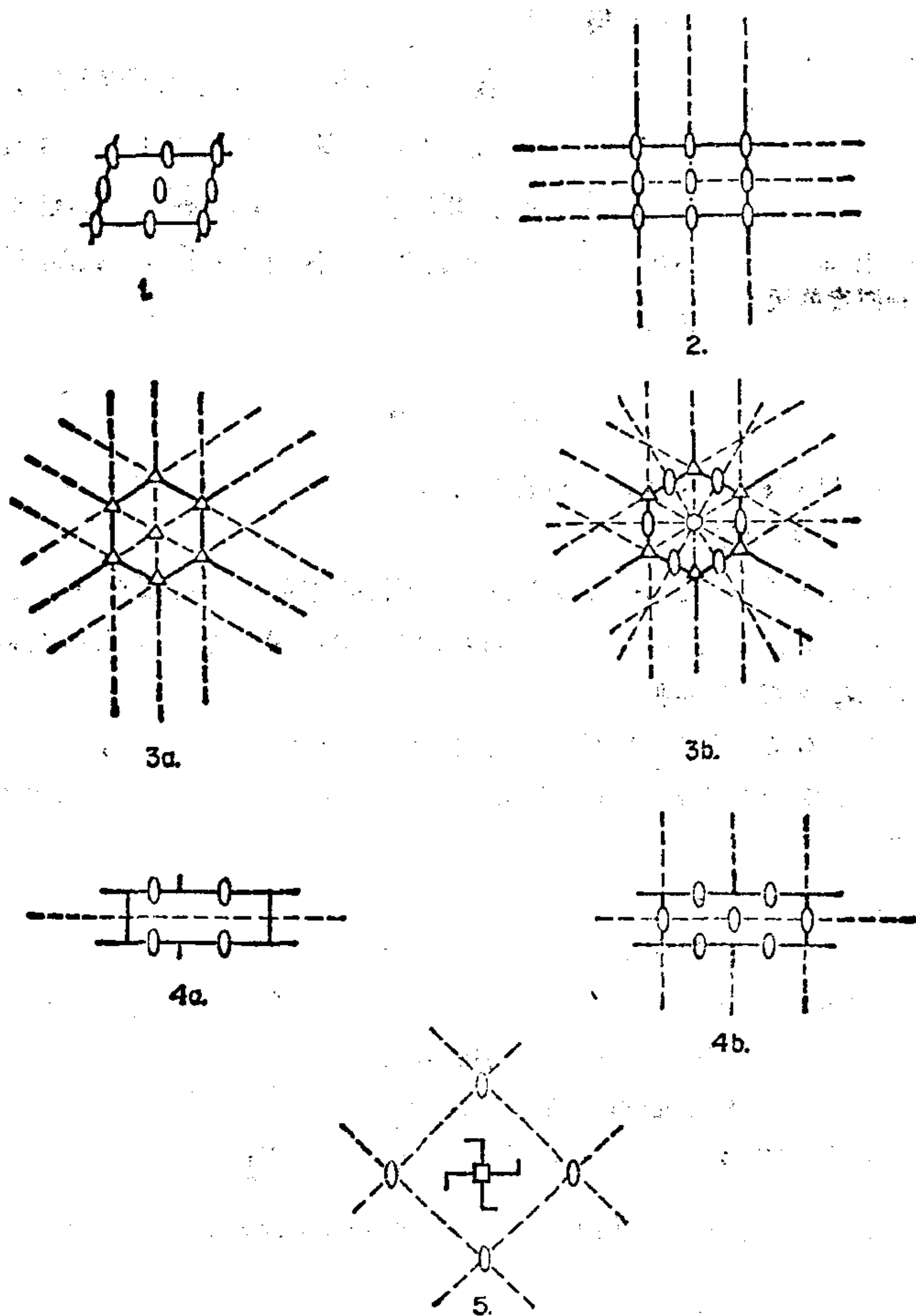


图 58 图案的反射对称与旋转对称

与反射对称就不用再作任何解释了.

可以再画一张表明这个图案的基本滑移反射的图，与上面表明旋转对称及反射对称的图一起可以说明这个图案的对称变换群中的全部元素. 一般可用观察法来画出基本滑移反射的轴（见练

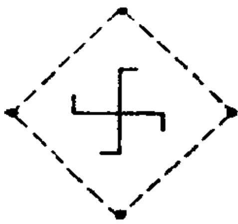


图 59　图案 5 的另一种图案单元

习 6 与 7).

最后，我们指出图案单元的选择方法不是唯一的．例如，图 59 画出了图案(5) 的另一种选择图案单元的方法．如果我们把图案旋转 45° 角，那末自然会用图 59 作为它的图案单元．

练　　习

1．验证图案 4 a 及 4 b 的对称变换群的元素 $T'_4G'_4$ 是基本滑移反射．画出这些滑移反射的轴．

2．验证 $G'_4T_4=T_4^{-1}G'_4$ 与 $G'_4T'_4=T'^{-1}_4G'_4$、描述基本滑移反射$(G'_4)^3$．

3．画出图案 5 的所有水平滑移轴及铅垂滑移轴．把关于这些轴的滑移反射写成 G_5 及 G'_5 的乘积．

4．求证 $G_5G'_5$ 是半周旋转．画出这个半周旋转的中心．

5．一个平面图案如果有二个滑移反射，它们的轴互相垂直，那末它也有旋转对称．

[提示：用练习 4．]

如果这二根轴任意相交，结论还成立吗？

6．分别就以下两种情况求一个滑移反射与一个半周旋转的乘积：

(i) 半周旋转中心在滑移反射的轴上；

(ii) 半周旋转中心不在滑移反射的轴上．

(图 57 的图案 4 是情况 (i)，图案 5 是情况(ii))

7．证明一个$\frac{\pi}{2}$角的旋转与一个关于水平轴的滑移反射的乘积是滑移反射，它的轴与水平轴成$\frac{\pi}{4}$角．

(图 57 中的图案 5 是本题的一例)

8．验证六边形图案 (3 b)的对称变换群包含一个滑移反射，它是$\frac{1}{2}$ $\mathbf{a}_1$ 的平移以及关于轴 L 的反射的乘积，L 是平分六边形相邻二边的水平方向的直线(图 60)．这个滑移反射不是基本滑移反射，试证明它可以写成乘积 T'_3S，使 T'_3 是 $\mathbf{a}'_3$ 决定的平移，S 是关于直线 C 的反射，直线 C 过六边形的

中心.

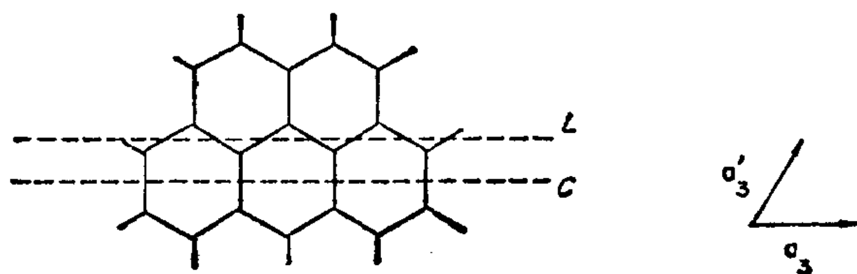


图 60 关于轴 L 的滑移反射

9. 一个图案的所有平移，滑移反射与恒等变换的集合是不是这个图案的对称变换群的子群?

10. 证明图 57 的图案 (1) 的对称变换群由平移 T_1 及 T_1' 生成. 并证明图案(2)的对称变换群由 T_2、T_2' 及二个反射生成——关于过图案单元中心 O 的水平直线与铅垂直线的反射 S_h 与 S_v.

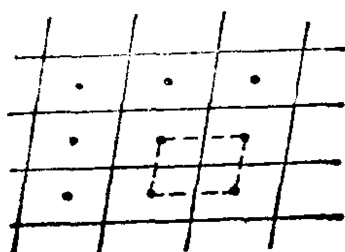
[提示：一个对称变换(等距变换)由图案单元及其象完全确定.对O点所在的图案单元使用 S_h, S_v, T_2, T_2' 这四个变换的乘积,便可知道这个图案单元可以通过对称变换映到任意一个图案单元. 同理可证，图 57 的其余几个图案的对称变换群至多由 4 个生成元生成.]

8.2 点阵

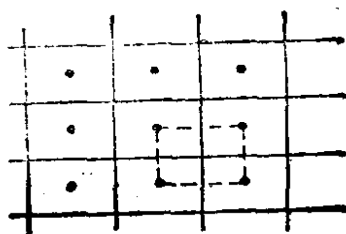
设$\mathscr{G}$是某个平面图案的对称变换群. 在这个平面图案的图案单元中任取一点. 在$\mathscr{G}$中的平移及滑移反射的作用下，这个点的像散布在整个平面上. 这样得到的点与图案单元一一对应. 这些点的全体所成的点集本身也是一个无限的平面图案. 有时我们可能选择单元代表点，不仅使$\mathscr{G}$中的平移与滑移反射是所得到的点集的对称变换，而且使$\mathscr{G}$中的旋转以及反射也是这一点集的对称变换. 我们就用前一节的例子来说明这种可能性(图 61).

在图 61 中，每种图案的单元代表点组成的点集是一个点阵. 在晶体学里，点阵是指点的一种排列方式，这种排列方式使每个点的周围情况都相同. 这里我们讨论平面点阵. 而若要对晶体的内

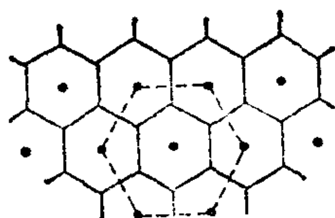
在结构作完整的描述，则空间点阵起着重要的作用.



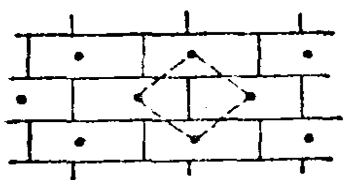
1.一般点阵



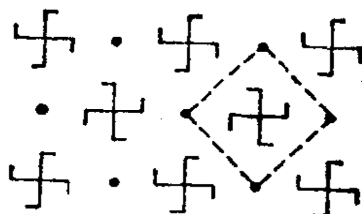
2. 矩形点阵



3. 六边形点阵



4. 菱形点阵



5. 正方形点阵

图 61　五类平面点阵

上图的点阵(2)，(3)，(4)，(5)都是点阵(1)的特例. 在一般点阵中，一个基本平行四边形(虚线所示)的边之间、角之间都没有任何特殊关系.

请注意在图 57 的图案(4a)中，由砖的中心点所成的点集不是一个点阵. 然而，如果我们在某个水平反射轴上取单元代表点，并使这一点与一个滑移反射轴的距离为砖长的四分之一，则我们得到一个菱形点阵，如点阵(4)所示. 如果一个图案的对称变换群包含基本滑移反射，则我们总可以找到这个图案的一个菱形点阵.

我们还要着重指出，一个图案的对称变换群一定是这个图案的点阵的对称变换群的一个子群．但是，有些特殊的图案的点阵的选取不是唯一的．对图案(4a)，右图的矩形点阵可以代替图61的菱形点阵．

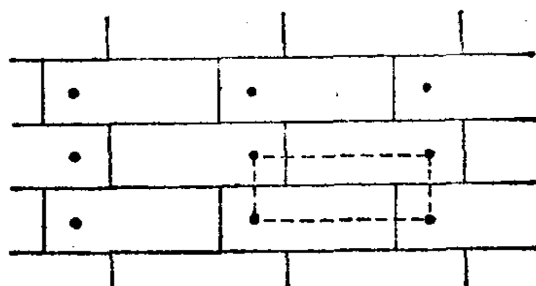


图 62　图案4a的矩形点阵

一个图案的点阵可以看成是一个骨架，图案建立在这个骨架上．构造图案的过程往往会减少对称．但是图案(1)，(2)，(3b)及(4b)的对称变换群等于它的点阵的对称变换群．

值得注意的是只可能有图61中画出的五类点阵．这个结论使人大吃一惊，难道正五边形的重复图案的点阵不是第六种点阵吗？如果你怀疑这一点，建议你动手去画一张正五边形的平面图案*．在下面一节中，我们将要证明平面上的一个重复图案的旋转对称只可能是旋转$\frac{\pi}{2}$或$\frac{\pi}{3}$的倍数角．利用这个晶体学关于旋转对称的制约，不难证明只可能有五类平面点阵．

练　　习

11．画出图案(4a)的各块砖的中心，验证它们不是点阵．

12．画一个由点构成的图案，使这图案是一个菱形点阵．验证这个点阵的对称变换群包含许多基本滑移反射，它们的轴是一组互相垂直的直线．

13．试说明一个菱形点阵可以看成一个带中心的矩形点阵，即一个矩形点阵连同各个矩形的中心所构成的图形．

8.3　晶体学中关于旋转对称的制约

对不同的平面对称变换群进行分类本来是一件艰难的工作，

* 事实上正五边形不可能布满整个平面，因此没有点阵——译者注．

但由于下面的定理，这个工作就比较容易一些了．这个定理对可能出现于这些群中的旋转对称加上一个很强的制约．

定理 I 平面上的重复图案的仅有的旋转对称是旋转$\frac{\pi}{2}$或$\frac{\pi}{3}$的倍数角．

证明 设这个图案有旋转θ角的旋转对称，θ不是$\frac{\pi}{2}$或$\frac{\pi}{3}$的倍数角．显然我们可以设

$$0<\theta<\pi.$$

令T是这个图案的移动距离最小的平移对称，它的位置向量是**a**．把**a**分别旋转θ及2θ角，得到向量**b**及**c**．这二个向量也给出了这个图案的二个平移对称，而且**a**、**b**、**c**这三个向量的和与差也给出这个图案的平移对称．图63(a)至(c)画出了可能发生的三种情况．在情况(a)：$\|\mathbf{b}-\mathbf{a}\|<\|\mathbf{a}\|$；在情况(b)：$\|\mathbf{a}+\mathbf{b}\|<\|\mathbf{a}\|$；在情况(c)：$\|\mathbf{a}+\mathbf{c}\|<\|\mathbf{a}\|$．所以在任何情况下，总可以找到这个图案的一个平移对称，使它的位移小于$\|\mathbf{a}\|$，与T的取法矛盾．定理证毕．

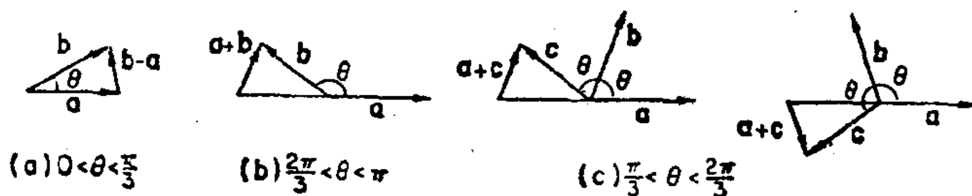


图 63 晶体学中关于旋转对称的制约的证明

关于十七种平面对称群的分类，建议读者参阅：

1. *International Tables for X-ray Crystallography*, Vol.1, 'Symmetry groups', Kynoch Press, 1965.

2. A.Bell & T.Fletcher, *Symmetry Groups*, Association of Teachers of Mathematics, Vine Street Chambers,

Nelson, Lancashire, 1964.

练　　习

14. 证明恰有五类平面点阵.

[提示：用晶体学关于旋转对称的制约，把可能有的类型归结成平行四边形点阵(半周旋转的旋转对称)，六边形点阵$\left(\frac{\pi}{3}\right.$角的旋转对称$\left.\right)$和矩形点阵，然后再证明从矩形点阵中可以得到二类点阵：正方形点阵与带中心的矩形(菱形)点阵.]

15. 证明在带中心的矩形点阵中，矩形的中心可以通过平移变到矩形的顶点上，但不能由一个旋转变到顶点上.

8.4　空间群和点群

现在我们可以进入讨论三维对称变换群的阶段，而不会有太多的困难. 这里我们只作简单的介绍并举几个例子加以说明. 在一些较深的课本里容易找到完整的讲解.

有两类重要的三维对称变换群——空间群与点群.

空间群是图案无限重复的、包含三个无关平移的对称变换群. 在刻划晶体的内部结构时会遇到这类群. 与二维时一样(见8.3节)，晶体学关于旋转对称的制约也适用于空间群. 一共有230个三维空间群. 这些群与由一些空间点阵构造出来的图形有关，这样的空间点阵共有十四种类型，称为布拉维斯点阵. 空间群超出了本书讨论范围，我们不进行讨论.

使空间中一个选定的点固定不动的等距变换组成的群称为点群. 当然，这样的群不包含平移. 从下面的定理可以清楚看出点群的重要性.

定理 II　一个有限物体的对称变换群$\mathscr{G}$是点群.

证明　用尽可能小的半径r的球S把这个物体包起来，使这个物体没有任何部分露在球的外面. 我们断言S被$\mathscr{G}$中的每个变

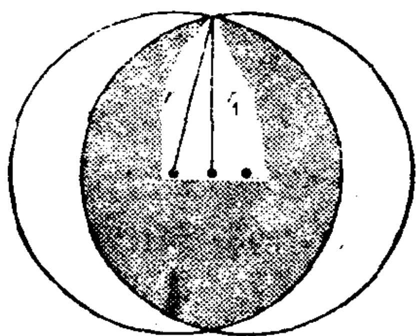


图 64 定理II的证明

换映到自身．否则，设$\mathscr{G}$中元素T将S映成另一个球S'．于是，这个物体也被S'包起来，因此它也被S与S'的“透镜形”公共部分包起来（图 64）．所以我们可以找到以r_1为半径的球把这个物体包起来，而r_1比r小．这与球S的取法矛盾．

我们知道，如果一个等距变换使球映到自身，那末必使球的中心固定不动．这样我们已经证明了$\mathscr{G}$的每个元素使S的中心固定不动．所以$\mathscr{G}$是点群．

存在无数多种不同的点群．这个结论是显然的，因为由周期是n的旋转C_n生成的群$\mathscr{C}_n(n=2,3,4,\cdots)$都是点群．在晶体学关于旋转对称的制约下，只有32个不同的点群，它们称为晶体点群．8.5 节给出了点群的几个例子．

练　习

16．证明把球映到自身的等距变换使球心固定不动．

8.5 点群的几个例子

例 1　循环群$\mathscr{C}_n$

循环群$\mathscr{C}_n$由关于n次轴的一个旋转C_n生成．许多物体的对称变换群都是循环群．四叶片的轮船螺旋桨的对称变换群是$\mathscr{C}_4$．过氧化氢的分子的对称变换群是$\mathscr{C}_2$．图 65 中把它的二次旋转轴画成铅垂方向的轴．我们作如下约定：如

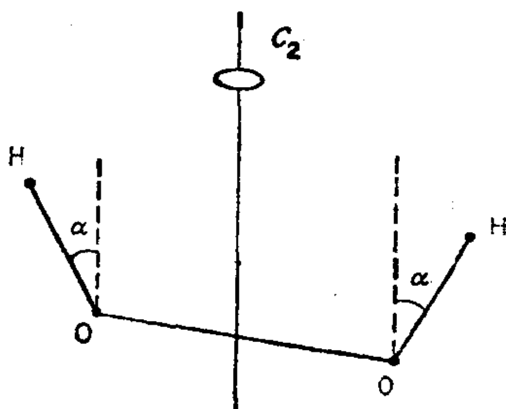


图 65　H_2O_2，对称$\mathscr{C}_2$

果在点群的许多旋转轴里，有一根轴的次数比其它任何轴的次数都要高，那末把这个轴看成是铅垂方向的轴．如果存在这样的轴，则称为主轴．

例 2　群$\mathscr{C}_{nh}$

$\mathscr{C}_{nh}$是由旋转C_n及反射σ_h生成，C_n是关于某根n次轴的旋转（总假设这根n次轴是铅垂方向的），σ_h是关于水平平面的反射．C_n与σ_h可交换，因此$\mathscr{C}_{nh}$是交换群．$\mathscr{C}_{nh}$有$2n$个元素，它们是$C_n^i\sigma_h^j(i=0,1,\cdots,n-1;\ j=0,1)$．群$\mathscr{C}_{1h}$由$\sigma_h$及恒等元这二个元素组成．$\mathscr{C}_{1h}$通常记成$\mathscr{C}_s$（$s$是德文“镜面”一词的第一个字母）．注意$\mathscr{C}_{nh}$可以写成直积

$$\mathscr{C}_{nh}=\mathscr{C}_n\times\mathscr{C}_s.$$

在一个等边三角形状的薄片的三个角上都挖去一个L形的小槽，所得到的薄片的对称变换群是$\mathscr{C}_{3h}$（如果不挖这些槽，原来的薄片的对称变换群比$\mathscr{C}_{3h}$更大一些）．平面分子反二氯乙稀的对称变换群是$\mathscr{C}_{2h}$（图 66）．

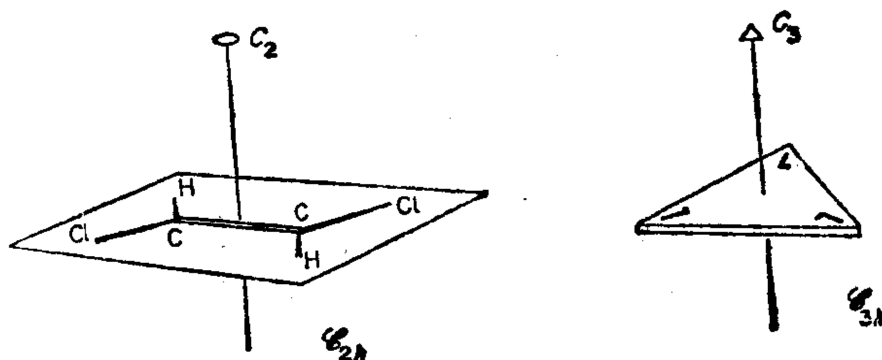


图 66　对称变换群$\mathscr{C}_{nh}$

例 3　群$\mathscr{C}_{nv}$

$\mathscr{C}_{nv}$由旋转C_n与反射σ_v生成，C_n是关于某n次主轴（假设铅垂方向）的旋转，σ_v是关于某个过主轴的铅垂方向的平面的反射．所以，共有n个铅垂方向的反射平面，它们分别是反射$\sigma^{(r)}=\sigma_vC_n^r(r=0,1,\cdots,n-1)$的反射平面．下面是群$\mathscr{C}_{nv}$的定义关系

式：

$$(\sigma_v)^2=1,\quad (C_n)^n=1,\quad C_n\sigma_v=\sigma_v C_n^{-1}. \qquad (1)$$

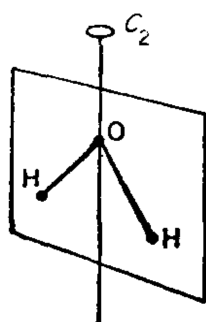


图 67 H_2O，对称变换群 $\mathscr{C}_{2v}$

与 7.1 节例 8 比较，群 $\mathscr{C}_{nh}$ 显然与 $2n$ 阶的平面二面体群同构．例 8 的平面 n 边形看成是三维空间中的水平平面上的图形时，它有一个铅垂的 n 次轴和 n 个铅垂的反射平面，这些平面与多边形相交于 $L, L'\cdots$.

水分子的对称变换群是 $\mathscr{C}_{2v}$(图 67)．氨分子的对称变换群是 $\mathscr{C}_{3v}$(6.1 节图 35)

例 4 三维的二面体群 $\mathscr{D}_n$．群 $\mathscr{D}_{nh}$．

三维的二面体群 $\mathscr{D}_n$ 只含有旋转．$\mathscr{D}_n$ 由旋转 C_n 及半周旋转 C_2 生成，C_n 是关于某 n 次主轴(假设铅垂)的旋转，C_2 是关于某(二重)水平轴的旋转．所以共有 n 根水平的二次轴，它们之间的夹角是 $\frac{2\pi}{n}$ 的整数倍．下面是 $\mathscr{D}_n$ 的定义关系式*：

$$(C_2)^2=1,\quad (C_n)^n=1,\quad C_nC_2=C_2C_n^{-1} \qquad (2)$$

从代数观点来看，$\mathscr{C}_{nv}$ 与 $\mathscr{D}_n$ 是同构的二个群．但是从几何观点来看，它们是不同的二个群．$\mathscr{D}_n$ 只包含旋转，而 $\mathscr{C}_{nv}$ 包含反射．

一个正六棱柱(见 7.1 节的图 52)的旋转对称变换所成的群是 $\mathscr{D}_6$．这个棱柱还有对称 σ_h，σ_h 是关于一个水平平面的反射，这个水平平面把棱柱分成二个相等的部分．σ_h 与 $\mathscr{D}_6$ 的所有元素都可以交换．这个棱柱的对称变换群是 $\mathscr{D}_{6h}=\mathscr{D}_6\times\mathscr{C}_s$，它是 $12\times 2=24$ 阶群．

在上述例题中，如果取 n 等于 2，3，6，可以得到 32 个晶体点群中的一些群．我们知道正方体的对称变换群是晶体点群，所以它的所有子群也都是晶体点群．四面体群是其中的一个子群．

---

* 如果 $n=2$，水平旋转的记号应与铅垂旋转的记号有所区别，例如可记成 $C_2^{(1)}$．

我们已在7.4节中讨论过四面体群.

最后我们要指出晶体学关于旋转对称的制约仅适用于晶体.而分子可能有$n\neq 2, 3, 4, 6$的n次对称主轴(但这种可能性极小). 例如,二茂铁$(C_5H_5)_2Fe$有一个五次轴.

进一步阅读可参见:

1. H. H. Jaffe & M. Orchin, *Spmmetry in Chemistry*, Wiley,1965.

2. R. McWeeny, *Symmetry, an Introduction to Group Theory and its Applications*, Pergamon Press,1963.

3. R,M.Hohstrasser, *Molecular Aspects of Symmetry*, W.A.Benjamin,1966.

4. H. S. M. Coxeter 和 W. O. J. Moser, *Generators and Relations for Discrete Groups*, Second Edition, Chapter 4, Springer,1965.

练　　习

17. 把六棱柱的铅垂方向的反射对称写成乘积$R\sigma_h$,使R在$\mathscr{D}_6$中.

18. 把六棱柱的两个底面都涂上两种颜色,如图68所示. 证明这个涂色的棱柱的旋转对称所成的群是$\mathscr{D}_3$,并且证明这个涂色棱柱有三个反射对称,三个伪旋转对称,它的全部对称成的对称变换群是12阶群,记成$\mathscr{D}_{3d}$.

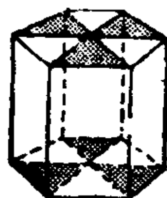


图 68　涂色六棱柱

19. 证明$\mathscr{D}_{3d}$是$\mathscr{D}_{6h}$的子群.

20. 求6.4节练习42的涂色正方体的对称变换群.

21. 证明6.4节练习43的正方体的对称变换群是循环群，它由伪旋转S_4生成. 这个群记成$\mathscr{S}_4$.

22. (i)证明正方体(不涂色的)的旋转对称的群是24阶群. 这个群叫做八面体群 $\mathscr{O}$. 这个名称的来源是因为只要把正八面体的六个顶点放在正方体各个面的中心位置处,就可以使正八面体内接于正方体. 所以正八面体与正方体的对称变换群是相同的.

(ii) 证明正方体的完全对称变换群是48阶的群. 记成 $\mathscr{O}_h$.

23. 求平面的、非线性分子NOCl的对称变换群.

24. 试画出下述分子的对称变换群的结构图:

(i) 氯甲烷 CH_3Cl,对称变换群 $\mathscr{C}_{3v}$.

(ii) 四氧化二氮 N_2O_4,对称变换群 $\mathscr{D}_{2h}$.

(iii) 苯 C_6H_6,对称变换群 $\mathscr{D}_{6h}$.

(iv) 六氟化铀 UF_6,对称变换群 $\mathscr{O}_h$.

第三部分

矩阵论及其应用

第九章　线性变换与矩阵

9.1　引论

我们已经看到变换在对称性问题中的重要性，并且用纯粹几何的方法学得了不少这方面的知识. 但若要求得到有关旋转、反射等的细节信息时，纯粹几何的方法有其局限性. 例如，设 C_3^y 为绕一条水平轴（y 轴）转 $\frac{2\pi}{3}$ 的旋转，而 C_3^z 为绕一条与之相交的铅垂轴（z 轴）转 $\frac{2\pi}{3}$ 的旋转，试求它们的积 $C_3^y C_3^z$（图 69）. 这个积是

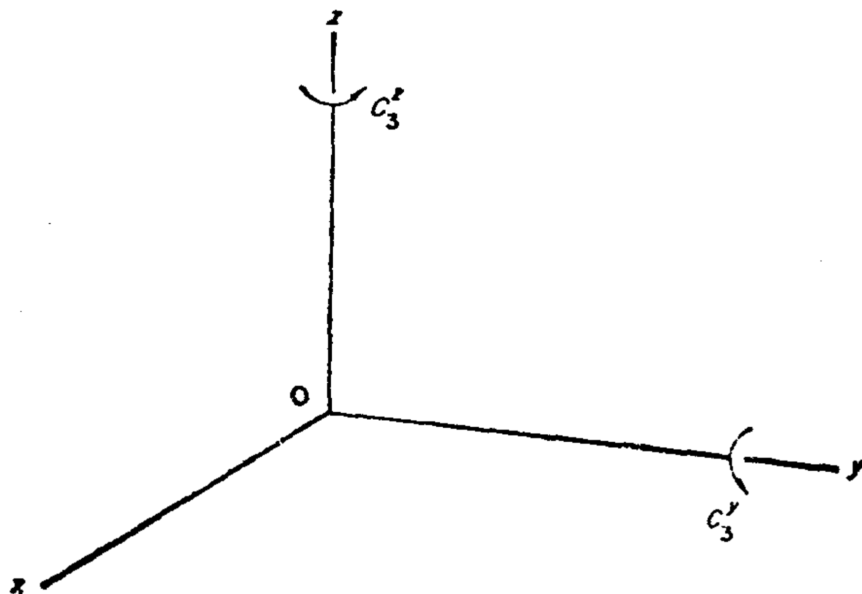


图 69　旋转问题

保持点 O 不动的正等距变换，从而是旋转（参看 6.4 节）. 你能求出它的旋转轴与旋转角吗？

我们将说明，如果在一个坐标系中把这些变换表示出来，求此问题的解可以归结为简单的数值计算. 当然，以一个坐标系作参

考的想法对我们来说已不是新的了．例如，在第四章中，我们定义向量的点积和叉积这两个重要概念时就这样做过．

在开始上述过程之前，我们必须介绍线性变换的概念．

9.2 线性变换

我们从一个简单的几何例子开始．设R是平面上绕点O沿逆时针方向的旋转．令α为其旋转角，$R(P)=P'$为P在R下的像．显然，R可以看成为位置向量的向量空间中由法则

$$R(\mathbf{p})=\mathbf{p}',\text{或简记为 } R\mathbf{p}=\mathbf{p}' \quad (1)$$

所确定的变换，此处$\mathbf{p}$与$\mathbf{p}'$分别是点P与P'的位置向量．这样处理后便有一系列重要的结果：由于把变换看成为在向量空间上的作用，就有可能在向量空间代数的基础上建立起变换的代数．

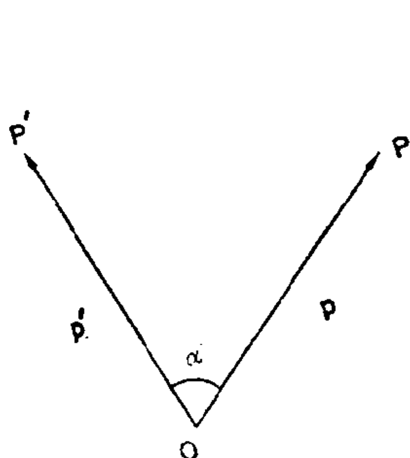


图 70　绕点O的旋转

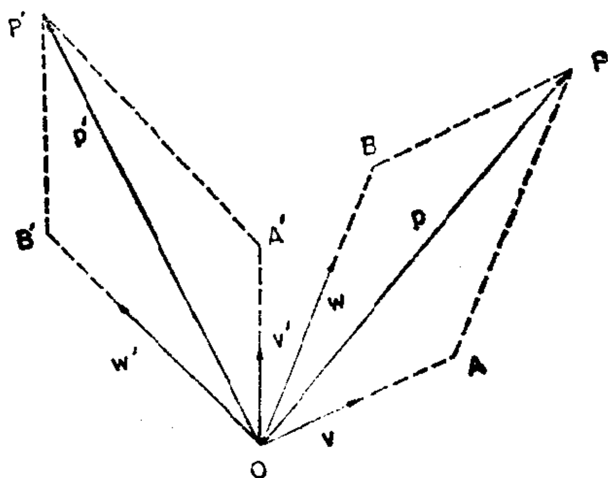


图 71　线性组合的旋转

考虑向量$\mathbf{v}$与$\mathbf{w}$的任意线性组合$\mathbf{p}=k\mathbf{v}+l\mathbf{w}$（在图71中，$k=2,l=\frac{3}{2}$）在$R$下的像．平行四边形$OAPB$被转成具有相同形状的平行四边形$OA'P'B'$．因而

$$\mathbf{p}'=(k\mathbf{v}+l\mathbf{w})'=k\mathbf{v}'+l\mathbf{w}', \quad (2)$$

或等价地写为

$$R(k\mathbf{v}+l\mathbf{w})=kR\mathbf{v}+lR\mathbf{w}. \tag{3}$$

这表明，两个向量 $\mathbf{v}$ 与 $\mathbf{w}$ 的一个线性组合在 R 下的像等于像向量 $R\mathbf{v}(=\mathbf{v}')$ 与 $R\mathbf{w}\ (=\mathbf{w}')$ 的同样的线性组合．我们简单地说 R 保持线性．向量空间的一个变换若保持线性，则称之为线性变换．

在数学中有许许多多线性变换的重要例子．最为人们所熟悉的例子之一是定义在所有可微函数组成的向量空间上的微分算子 $\frac{d}{dx}$．保持一点 O 不动的二维或三维的任一等距变换，都是关于点 O 的位置向量的向量空间上的线性变换．对于平面的等距变换，这一结论可用像图 71 那样的图形来证明：保持点 O 不动的任意等距变换把平行四边形 $OAPB$ 映到具有相同形状的平行四边形 $OA'P'B'$ 上，因此，这个等距变换保持线性．

练　　习

1．证明，向量空间的线性变换把零向量 $\mathbf{0}$ 映到零向量．

2．设 R 是线性变换且 $R\mathbf{v}=\mathbf{v}'$，证明 $R(-\mathbf{v})=-\mathbf{v}'$．

3．设 S 是关于直线 L 的平面反射．画一草图说明 S 是关于 L 上一个原点的位置向量的向量空间上的线性变换．

4．若 S 为到 L 上的正交射影，重复练习 3．

5．设 R 与 S 是向量空间 $\mathscr{V}$ 上的线性变换，证明它们的积 RS 也是线性变换．

6．证明，平面上的平移不是线性变换．

7．画一草图说明，三维空间中绕一条固定轴 L 的旋转是关于 L 上一个原点的位置向量的向量空间上的线性变换．

9.3　线性变换的矩阵

再考虑上节提到的旋转 R．我们引进一个以 O 为原点的直角坐标系（图 72(a)）．图上把旋转角 α 画成了锐角，但下面的讨论对任意的角 α 都适用．点 $P\ (x_1,\ x_2)$ 在 R 下的像记为

$P'(x_1', x_2')$. 为了对此坐标系定义出 R，我们必须求出用点 P 的坐标 x_1, x_2 表示点 P' 的坐标 x_1', x_2' 的公式.

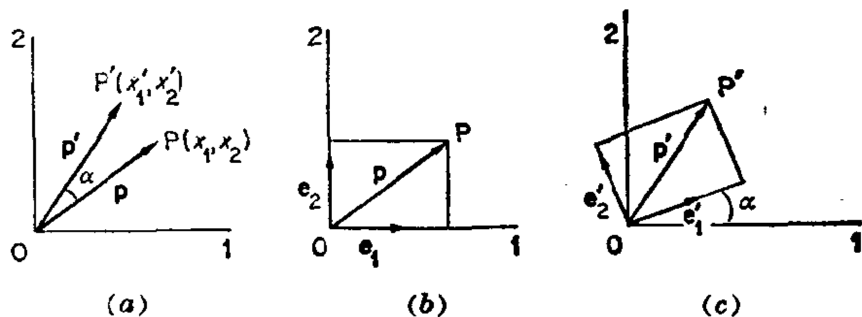


图 72 在一个坐标系中的旋转

设 $\mathbf{e}_1, \mathbf{e}_2$ 为坐标轴方向上的单位向量，则

$$\mathbf{p} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2, \tag{4}$$

而由于 R 是线性变换，有

$$\mathbf{p}' = x_1\mathbf{e}_1' + x_2\mathbf{e}_2' \tag{5}$$

(参看图 72(b)与(c)). 单位向量 $\mathbf{e}_1'$ 与 $\mathbf{e}_2'$ 在坐标系中的分量容易求出：

$$\mathbf{e}_1' = (\cos\alpha,\ \sin\alpha), \qquad \mathbf{e}_2' = (-\sin\alpha,\ \cos\alpha).$$

所以，公式(5)变成

$$\begin{aligned}\mathbf{p}' = (x_1', x_2') &= x_1(\cos\alpha, \sin\alpha) + x_2(-\sin\alpha, \cos\alpha)\\ &= (x_1\cos\alpha - x_2\sin\alpha,\ x_1\sin\alpha + x_2\cos\alpha).\end{aligned}$$

从而

$$\left.\begin{aligned} x_1' &= x_1\cos\alpha - x_2\sin\alpha\\ x_2' &= x_1\sin\alpha + x_2\cos\alpha.\end{aligned}\right\} \tag{6}$$

在上述坐标系中，此公式定义了 R，因为它给出了任意一点 $P(x_1, x_2)$ 的像的坐标. 我们称公式(6)为在所给的坐标系中变换 R 的法则. 这个法则是由线性关系

$$x_1' = a_{11}x_1 + a_{12}x_2, \qquad x_2' = a_{21}x_1 + a_{22}x_2$$

定义的，这里

$$a_{11} = \cos\alpha, \quad a_{12} = -\sin\alpha, \quad a_{21} = \sin\alpha, \quad a_{22} = \cos\alpha.$$

我们称系数的阵列

$$\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix} \tag{7}$$

为在所给坐标系中R的矩阵.

注意到$e_1=(1,0)$在R下的像e_1'可由在公式(6)中置$x_1=1$, $x_2=0$而得到. 因此, e_1'的分量在R的矩阵的第一列中出现. 同样, e_2'的分量在第二列中出现.

用上述方法, 我们可以得到任意一个线性变换在一个特定的坐标系中的矩阵与线性法则.

例 1　反射的矩阵

求一个反射的矩阵, 设反射轴为与 01 轴倾角为β的直线L (图 73).

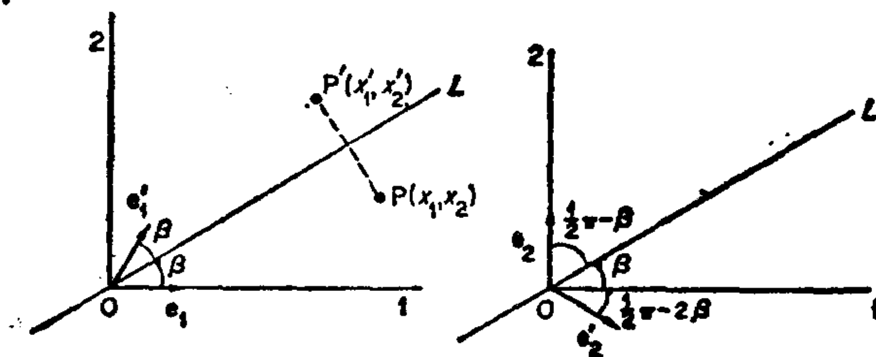


图 73　关于L的反射

矩阵的两列元素分别为e_1'与e_2'的分量, 所以我们得到该反射的矩阵及对应的线性法则如下:

矩阵: $\begin{pmatrix}\cos 2\beta & \sin 2\beta\\ \sin 2\beta & -\cos 2\beta\end{pmatrix}$,

法则: $$\begin{aligned}x_1'&=(\cos 2\beta)x_1+(\sin 2\beta)x_2\\ x_2'&=(\sin 2\beta)x_1-(\cos 2\beta)x_2.\end{aligned}$$

这里, e_2'的第一分量为$\cos\left(\frac{1}{2}\pi-2\beta\right)=\sin 2\beta$, 其第二分量

类似地计算．我们以 β 为锐角的情况作图示．当 β 为钝角时，同样的矩阵仍然适用．

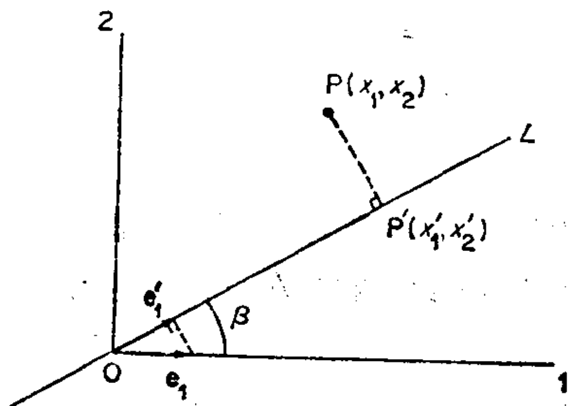


图 74　到 L 上的正交射影

例 2　正交射影的矩阵

在 9.2 节的练习 4 中我们曾看到，到一条直线 L 上的正交射影是关于 L 上一个原点的位置向量的向量空间上的线性变换．在所画的以 O 为原点的坐标系(图74)中，求该射影的矩阵和法则．

e_1' 是长度为 $\cos\beta$ 的向量，因此它的分量为 $\cos\beta\cdot\cos\beta$ 与 $\cos\beta\cdot\sin\beta$．对 e_2' 同样地计算，便得到如下结果：

矩阵：$\begin{pmatrix} \cos^2\beta & \cos\beta\ \sin\beta \\ \cos\beta\ \sin\beta & \sin^2\beta \end{pmatrix}$；

法则：$\begin{aligned} x_1' &= (\cos^2\beta)x_1 + (\cos\beta\ \sin\beta)x_2 \\ x_2' &= (\cos\beta\ \sin\beta)x_1 + (\sin^2\beta)x_2. \end{aligned}$

我们以 β 为锐角的情况作图示．当 β 为钝角时，同样的矩阵仍然适用．

若置 $\beta=0$，便得到了到水平轴(01 轴)上的正交射影的矩阵：

矩阵：$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$；　　法则：$\begin{aligned} x_1' &= x_1 \\ x_2' &= 0. \end{aligned}$

又，当 $\beta=\dfrac{\pi}{4}$ 时，我们得到了到与两轴倾斜成等角的直线上的正交射影的矩阵：

矩阵：$\begin{pmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}$；　法则：$\begin{aligned} x_1' &= \frac{1}{2}x_1 + \frac{1}{2}x_2 \\ x_2' &= \frac{1}{2}x_1 + \frac{1}{2}x_2. \end{aligned}$

例 3 下列矩阵各表示平面上的什么线性变换（在通常的直角坐标系中）？

(i) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; (iii) $\begin{pmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{pmatrix}$;

(iv) $\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$; (v) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; (vi) $\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$.

(i) 顺时针旋转$\frac{\pi}{2}$$\left(在矩阵(7)中置 \alpha=-\frac{\pi}{2}\right)$.

(ii) 关于 02 轴的反射$\left(在例 1 中置 \beta=\frac{\pi}{2}\right)$.

(iii) 到通过原点而与 01 轴成$\frac{\pi}{3}$角的直线上的正交射影.

(iv) 逆时针旋转$\frac{\pi}{3}$.

(v) 恒等变换 $x_1'=x_1$, $x_2'=x_2$.

(vi) 这是平行于 01 轴的切变的矩阵，其变换法则为

$$x_1' = x_1 + \frac{1}{2}x_2,$$
$$x_2' = \qquad x_2.$$

每一点在 01 轴方向上移动一段距离，这段矩离与该点在 01 轴上方的高度成正比（图 75）．切变角为 α，此处 $\tan\alpha=\frac{1}{2}$．请用一副扑克牌作切变变换的演示！

例 4 三维空间的旋转

现在要写出图 69 所示的坐标系中的旋转 C_3^y 与 C_3^z (9.1 节) 的矩阵与法则已是件容易的事了．

我们先考虑 C_3^y．用 $\mathbf{e}_1$,$\mathbf{e}_2$与 $\mathbf{e}_3$分别表示 Ox,Oy 与 Oz 轴上的

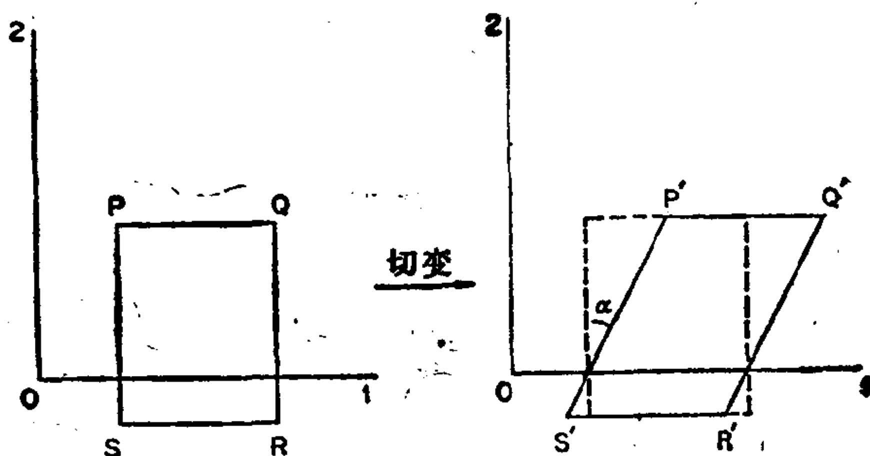


图 75 切变变换

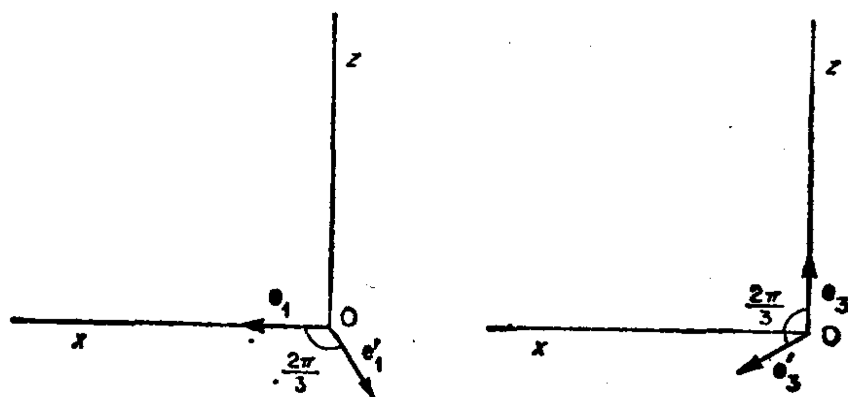


图 76 旋转问题中的 xz 平面

单位向量．显然 $C_3^y e_2 = e_2$，而 e_1 与 e_3 在 C_3^y 下的像如图 76 所示．据此，我们计算像向量的分量而得到

$$e_1' = \left(-\frac{1}{2}, 0, -\frac{\sqrt{3}}{2}\right), \ e_2' = (0,1,0), \ e_3' = \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right).$$

把行向量变成矩阵的列，我们便得到在此坐标系中表示 C_3^y 的矩阵．它是

$$A = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \end{pmatrix}.$$

因此，点 $P(x_1,x_2,x_3)$经过此旋转后，它的像是 $P'(x_1',x_2',x_3')$，其中

$$
\begin{aligned}
x_1' &= -\frac{1}{2}x_1 \qquad + \frac{\sqrt{3}}{2}x_3, \\
x_2' &= \qquad\qquad x_2, \\
x_3' &= -\frac{\sqrt{3}}{2}x_1 \qquad -\frac{1}{2}x_3.
\end{aligned}
$$

类似地求出 C_3^2 的矩阵为

$$
B=\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

$\mathscr{R}^n$ 的线性变换

设 T 是 $\mathscr{R}^n$ 的一个变换，把 $\mathbf{x}=(x_1,\cdots,x_n)$ 映到 $T\mathbf{x}=\mathbf{x}'=(x_1',\cdots,x_n')$。那么，$T$ 是线性变换当且仅当它满足形如

$$
\left.\begin{aligned}
x_1' &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
x_2' &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_n' &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n
\end{aligned}\right\} \tag{8}
$$

的法则．系数阵列

$$
\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}
$$

称为(8)式所定义的变换的矩阵．

这就推广了我们在二维和三维空间中参考一个坐标系所得出的结论，其证明留作练习．

练　　习

8．对钝角 α 验证旋转矩阵(7)．证明，在(7)中若以 $-\beta$ 代替 α，得到的矩阵对应于顺时针转 β 角的旋转．

9．对各种不同的角 β 验证反射的矩阵．

10．下列矩阵各表示平面上的什么线性变换（在通常的直角坐标系中）？

(i) $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$　(ii) $\begin{pmatrix}3 & 0\\ 0 & 3\end{pmatrix}$　(iii) $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$

(iv) $\begin{pmatrix}\frac{3}{5} & -\frac{4}{5}\\ \frac{4}{5} & \frac{3}{5}\end{pmatrix}$　(v) $\begin{pmatrix}\frac{3}{5} & \frac{4}{5}\\ -\frac{4}{5} & \frac{3}{5}\end{pmatrix}$　(vi) $\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix}$

(vii) $\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$　(viii) $\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$

11．试判断，以(3,2)，(−1,2)，(−1,−1)，(3,−1)为顶点的矩形在下列矩阵所表示的每个线性变换作用下各将发生什么情况？

(i) $\begin{pmatrix}3 & 0\\ 0 & 3\end{pmatrix}$（膨胀）；　(ii) $\begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix}$（收缩）；

(iii) $\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$（切变）；　(iv) $\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}$（正交射影）．

12．验证例4中 C_1^* 的矩阵如正文所述．

13．描述在直角坐标系中由下列矩阵所表示的线性变换：

(i) $\begin{pmatrix}\cos\alpha & -\sin\alpha & 0\\ \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix}$；　(ii) $\begin{pmatrix}\cos\alpha & 0 & -\sin\alpha\\ 0 & 1 & 0\\ \sin\alpha & 0 & \cos\alpha\end{pmatrix}$；

(iii) $\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}$．

14. 证明，如果 T 是 $\mathscr{R}^n$ 的线性变换，则 T 是由形如(8)的线性法则所定义的。

[提示：设

$\mathbf{e}_1=(1,0,0,\cdots,0)$，$\mathbf{e}_2=(0,1,0,\cdots,0)$，等等，

再假定

$T\mathbf{e}_1=(a_{11},a_{21},\cdots,a_{n1})$，$T\mathbf{e}_2=(a_{12},a_{22},\cdots,a_{n2})$，等等.

$\mathscr{R}^n$ 的任意向量 $\mathbf{x}$ 可以表为

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\cdots+x_n\mathbf{e}_n.$$

利用 T 的线性性质去计算 $\mathbf{x}'=T\mathbf{x}$.]

证明逆命题：如果 T 由线性法则(8)所定义，则 T 是 $\mathscr{R}^n$ 的线性变换.

9.4 两个正方矩阵的乘积

设 S 与 T 都是 $\mathscr{R}^n$ 的线性变换.（例如，当 $n=3$ 时，在一个坐标系中它们可能表示旋转或反射等.）设 S 的矩阵为 $A=(a_{ij})$，T 的矩阵为 $B=(b_{ij})$. 它们都是 $n\times n$ 矩阵. 我们将要定义矩阵的乘法使得线性变换 ST 的矩阵是 AB. 我们只考虑 $n=3$ 的情况，由此再推广到任意的 n 则是很容易的事了.

设 $\mathbf{x}=(x_1,x_2,x_3)$ 为 $\mathscr{R}^3$ 的任意向量. 置 $T\mathbf{x}=\mathbf{x}'$，$S\mathbf{x}'=\mathbf{x}''$. 则

$$(ST)\mathbf{x}=S(T\mathbf{x})=S\mathbf{x}'=\mathbf{x}''. \tag{9}$$

此式实际上已给出了 ST 的法则. 我们只要把(9)式用向量的分量表示出来就可以了.

按照法则

$$\left.\begin{aligned} x_1'&=b_{11}x_1+b_{12}x_2+b_{13}x_3,\\ x_2'&=b_{21}x_1+b_{22}x_2+b_{23}x_3,\\ x_3'&=b_{31}x_1+b_{32}x_2+b_{33}x_3 \end{aligned}\right\} \tag{10}$$

T 把 $\mathbf{x}$ 映到 $\mathbf{x}'$；然后再按照法则

$$\left.\begin{aligned}x''_1&=a_{11}x'_1+a_{12}x'_2+a_{13}x'_3,\\x''_2&=a_{21}x'_1+a_{22}x'_2+a_{23}x'_3,\\x''_3&=a_{31}x'_1+a_{32}x'_2+a_{33}x'_3\end{aligned}\right\}\qquad(11)$$

S 把 $\mathbf{x}'$ 映到 $\mathbf{x}''$. 把(10)式中 x'_1, x'_2 与 x'_3 的表达式代入(11)式,我们便得到 ST 的法则,它给出了 $\mathbf{x}$ 与它在 ST 下的像 $\mathbf{x}''$ 之间的联系:

$$\begin{aligned}x''_1&=(a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31})x_1+(a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32})x_2\\&\quad+(a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33})x_3\\&=\left(\sum_{k=1}^{3}a_{1k}b_{k1}\right)x_1+\left(\sum_{k=1}^{3}a_{1k}b_{k2}\right)x_2+\left(\sum_{k=1}^{3}a_{1k}b_{k3}\right)x_3,\\x''_2&=\left(\sum_{k=1}^{3}a_{2k}b_{k1}\right)x_1+\left(\sum_{k=1}^{3}a_{2k}b_{k2}\right)x_2+\left(\sum_{k=1}^{3}a_{2k}b_{k3}\right)x_3,\\x''_3&=\left(\sum_{k=1}^{3}a_{3k}b_{k1}\right)x_1+\left(\sum_{k=1}^{3}a_{3k}b_{k2}\right)x_2+\left(\sum_{k=1}^{3}a_{3k}b_{k3}\right)x_3.\end{aligned}$$

我们用规则

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\end{pmatrix}=\begin{pmatrix}c_{11}&c_{12}&c_{13}\\c_{21}&c_{22}&c_{23}\\c_{31}&c_{32}&c_{33}\end{pmatrix}$$

来定义矩阵的乘积 AB,这里

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+a_{i3}b_{3j}=\sum_{k=1}^{3}a_{ik}b_{kj}.$$

按此方法定义的矩阵乘法,矩阵 AB 正好表示线性变换 ST 在所给的坐标系中的矩阵.

有一个容易的方法记忆如何把两个矩阵相乘. 例如, 考察乘积 AB 的第二行与第三列交叉点上的元素 c_{23}. 我们有

$$c_{23}=a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33},$$

亦即,c_{23} 是由把 A 的第二行上的每个元素分别乘以 B 的第三列上

的对应元素，然后再把所得到的积加起来而得到的．如下所示：

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{pmatrix}\begin{pmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{pmatrix}=\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{pmatrix}.$$

注意这与两个向量的点积的求法十分相似．

推广之，我们便得到 $n\times n$ 矩阵乘法的定义．

定义 I 设 $A=(a_{ij}),B=(b_{ij})$是两个 $n\times n$ 矩阵，则它们的积 AB 是 $n\times n$ 矩阵(c_{ij})，其中第 i 行与第 j 列交叉点上的元素 c_{ij} 是由 A 的第 i 行上的每一个元素分别乘以 B 的第 j 列上的对应元素，并把所得到的积加起来而得到的：

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj} \tag{13}$$

矩阵的乘法是结合的，因为我们已经知道对应的变换积是结合的(6.2 节)．从而，三个或更多个 $n\times n$ 矩阵的乘积中，括号都可以省去，也可以作矩阵的幂 A^2,A^3 等等，均不会有歧义．

请注意，矩阵的乘法仅仅是对同样大小的正方矩阵定义的，还可能把这个定义推广到某些长方矩阵(9.5 节)．但是，若 $m\neq n$，一个 $m\times m$ 的矩阵绝不可能同一个 $n\times n$ 矩阵相乘．

例 5 求矩阵

$$A=\begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} \quad 与 \quad B=\begin{pmatrix} 2 & -2 \\ 1 & 0 \end{pmatrix}$$

的乘积．

$$\begin{aligned} AB&=\begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ 1 & 0 \end{pmatrix} \\ &=\begin{pmatrix} 1\cdot 2+(-2)\cdot 1 & 1\cdot(-2)+(-2)\cdot 0 \\ 0\cdot 2+3\cdot 1 & 0\cdot(-2)+3\cdot 0 \end{pmatrix} \\ &=\begin{pmatrix} 0 & -2 \\ 3 & 0 \end{pmatrix}. \end{aligned}$$

例 6 把 9.3 节例 4 中两个旋转矩阵相乘．

$$AB=\begin{pmatrix} -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} \frac{1}{4}+0+0 & \frac{\sqrt{3}}{4}+0+0 & 0+0+\frac{\sqrt{3}}{2} \\ 0+\frac{\sqrt{3}}{2}+0 & 0-\frac{1}{2}+0 & 0+0+0 \\ \frac{\sqrt{3}}{4}+0+0 & \frac{3}{4}+0+0 & 0+0-\frac{1}{2} \end{pmatrix}$$

$$=\begin{pmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{\sqrt{3}}{4} & \frac{3}{4} & -\frac{1}{2} \end{pmatrix}.$$

这就是本章引论中提及的变换 $C_3^y C_3^z$ 的矩阵．我们知道它表示一个旋转．在 9.6 节中我们将看到怎样从此矩阵中求出旋转轴与旋转角．

练　　习

15．求乘积 AB 与 BA，其中

$$\text{(i)}\ A=\begin{pmatrix} 2 & -6 \\ -1 & 3 \end{pmatrix},\ B=\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix};$$

$$\text{(ii)}\ A=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},\ B=\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}.$$

16．从练习 15(ii)推导 $\cos(\alpha+\beta)$与 $\sin(\alpha+\beta)$的公式．

17．用相乘对应矩阵的方法证明旋转 C_3^y 与 C_3^z 是不可交换的．

18．设

$$A=\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}.$$

证明，对所有正整数 n 都有 $A^n=A$。这一结果的几何解释是什么？

19．找一个 2×2 矩阵 I 使得对所有 2×2 矩阵 A 都有 $AI=A$。

20．用矩阵的方法证明，关于直线 $x_2=(\tan\alpha)x_1$ 的反射接上关于直线 $x_2=(\tan\beta)x_1$ 的反射，等价于一个旋转。试求其旋转角。

9.5 长方矩阵

上一节中我们定义了两个 $n\times n$ 矩阵（或称正方矩阵）的乘法．对于许多目的而言，考虑长方矩阵的乘法也是很有用处的．长方矩阵也就是行数与列数不一定相等的矩阵．例如，考虑矩阵

$$A=\begin{pmatrix}1 & 2 & 0\\ -3 & 1 & 1\end{pmatrix} \quad 与 \quad B=\begin{pmatrix}3 & 0 & 1 & 1\\ 5 & 0 & -1 & 1\\ 1 & 0 & 6 & 1\end{pmatrix}.$$

我们用同正方矩阵相乘一样的规则来作出矩阵的乘积 $AB=C$：矩阵 C 中的元素 c_{ij} 是由矩阵 A 的第 i 行的每一个元素乘以矩阵 B 的第 j 列的对应元素，然后再把所得到的积相加而得到的．在我们的例子中，元素 c_{23} 是这样得到的：

$$\begin{pmatrix}\cdot & \cdot & \cdot\\ -3 & 1 & 1\end{pmatrix}\begin{pmatrix}\cdot & \cdot & 1 & \cdot\\ \cdot & \cdot & -1 & \cdot\\ \cdot & \cdot & 6 & \cdot\end{pmatrix}=\begin{pmatrix}\cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & 2 & \cdot\end{pmatrix}.$$

矩阵 A 乘以矩阵 B 的这个规则只有当 A 的列数与 B 的行数相同时才可使用．

定义 II 一个 $m\times n$ 矩阵 A 与一个 $r\times p$ 矩阵 B 的积 AB 只有当 $n=r$ 时才有定义．当这个条件满足时，AB 是由按上文所说的办法把 A 的行与 B 的列"相乘"而得到的 $m\times p$ 矩阵．

在我们的数值实例中，A 为 2×3 矩阵，B 为 3×4 矩阵，于是

AB 为 2×4 矩阵．完整的算式是

$$\begin{pmatrix} 1 & 2 & 0 \\ -3 & 1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 1 & 1 \\ 5 & 1 & -1 & 1 \\ 1 & 0 & 6 & 1 \end{pmatrix}=\begin{pmatrix} 13 & 2 & -1 & 3 \\ -3 & 1 & 2 & -1 \end{pmatrix}.$$

注意，乘积 BA 是没有意义的．

例 7 3×3 矩阵与 3×1 矩阵的乘积是 3×1 矩阵，亦即一个单列．计算

$$\begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 3 \\ 2 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}.$$

这个乘积等于

$$\begin{pmatrix} 2\cdot 1+0\cdot(-3)+(-1)\cdot 1 \\ 1\cdot 1+1\cdot(-3)+ \quad 3\cdot 1 \\ 2\cdot 1+1\cdot(-3)+ \quad 1\cdot 1 \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

练　　习

21．设

$$A=\begin{pmatrix} 1 & -2 \\ 2 & -4 \\ -3 & 6 \end{pmatrix}, \qquad B=\begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

此时 AB 与 BA 都有定义。求这些乘积。

22．设

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

求

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ 以及 } A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

23．设 A 与 B 如练习 21，再设

$$C=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

证明$(AB)C=A(BC)$。可以一般性地证明，长方矩阵的乘积是结合的. 顺便提一句，这里$(BA)C$是没有定义的.

24. 设A是$m\times n$矩阵. 证明，A^2有定义当且仅当$m=n$.

9.6 关于列向量的约定

设T为$\mathscr{R}^3$的线性变换，其法则定义为

$$\left.\begin{aligned} x_1'&=a_{11}x_1+a_{12}x_2+a_{13}x_3 \\ x_2'&=a_{21}x_1+a_{22}x_2+a_{23}x_3 \\ x_3'&=a_{31}x_1+a_{32}x_2+a_{33}x_3. \end{aligned}\right\} \tag{14}$$

这个法则可以用矩阵的形式方便地写成

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{15}$$

这种写法与矩阵的乘法规则是一致的.

现在我们约定，把$\mathscr{R}^3$的向量写成3×1矩阵，或者说，写成列向量：$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. 在此约定下，$T$的法则写成简单的形式

$$\mathbf{x}'=A\mathbf{x}, \tag{16}$$

这里A为T的矩阵.

同样，$\mathscr{R}^n$的线性变换T也可以写成(16)的形式，其中$\mathbf{x}$与$\mathbf{x}'$是$\mathscr{R}^n$的列向量，而A是T的$n\times n$矩阵. 我们把A与T等同起来，从而把A本身看成$\mathscr{R}^n$的线性变换，这样做是很方便的.

例 8 在平面上给定通常的直角坐标系，试求点$P(-\sqrt{3},-1)$经过绕原点逆时针转$\frac{\pi}{3}$的旋转作用后的像P'.

该旋转的矩阵为

$$A=\begin{pmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}$$

(参看 9.3 节例 3 (iv)). 点 P 的位置向量写成列向量是

$$\mathbf{p}=\begin{pmatrix}-\sqrt{3}\\ -1\end{pmatrix},$$

这个旋转把向量 $\mathbf{p}$ 映到

$$\mathbf{p}'=A\mathbf{p}=\begin{pmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}\begin{pmatrix}-\sqrt{3}\\ -1\end{pmatrix}=\begin{pmatrix}0\\ -2\end{pmatrix}.$$

因此,点 P 的像为 $P'(0,-2)$. 画一草图验证这个结果.

例 9 旋转问题的解

我们回到解释旋转 $C_3^y C_3^z$ 的矩阵的问题上来. 在 9.4 节例 6 中,我们已经计算出这个矩阵,它是

$$R=\begin{pmatrix}\frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0\\ \frac{\sqrt{3}}{4} & \frac{3}{4} & -\frac{1}{2}\end{pmatrix}.$$

为了描述 R 所表示的旋转,我们必须找到旋转轴(到现在为止我们仅知道它通过坐标系的原点 O), 并且求出旋转角. 先求旋转轴. 设

$$\mathbf{v}=\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}\neq\mathbf{0}$$

是轴上异于O的某一点的位置向量．该旋转保持$\mathbf{v}$不动，亦即

$$R\mathbf{v}=\mathbf{v}. \tag{17}$$

从方程(17)，我们得到(齐次)线性方程组

$$\left.\begin{aligned}\frac{1}{4}x_1+\frac{\sqrt{3}}{4}x_2+\frac{\sqrt{3}}{2}x_3&=x_1\\ \frac{\sqrt{3}}{2}x_1-\frac{1}{2}x_2\qquad\qquad&=x_2\\ \frac{\sqrt{3}}{4}x_1+\frac{\sqrt{3}}{4}x_2-\frac{1}{2}x_3&=x_3,\end{aligned}\right\} \tag{18}$$

它的通解是$x_2=x_3=k$，$x_1=\sqrt{3}k$(k任意)．因此，旋转轴由所有位置向量形如$k\mathbf{v}$的点所组成，这里

$$\mathbf{v}=\begin{pmatrix}\sqrt{3}\\1\\1\end{pmatrix}.$$

要确定旋转角，我们找一个垂直于$\mathbf{v}$的向量$\mathbf{w}$，并且计算它被这个旋转所转过的角度(图 77)．

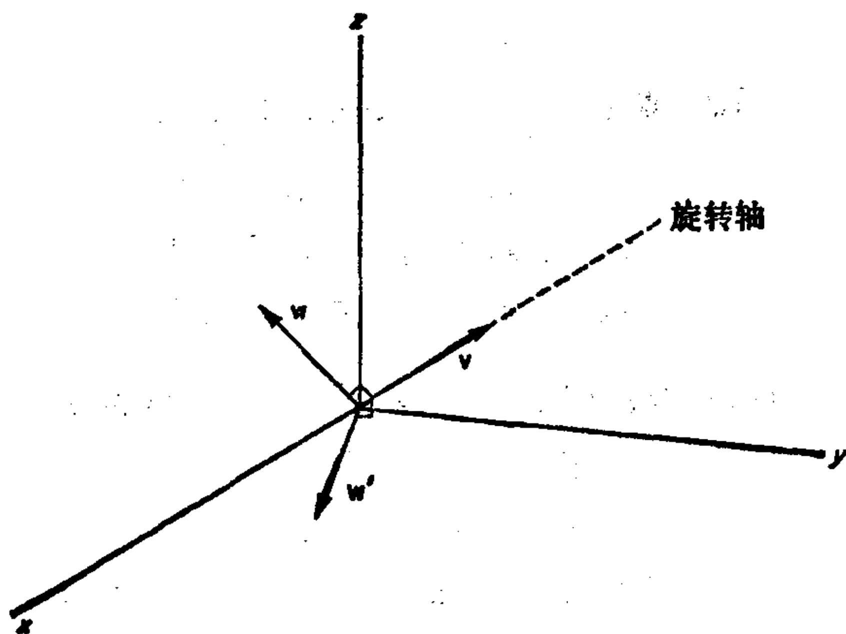


图 77

一个适当的选法是

$$\mathbf{w}=\begin{pmatrix}0\\-1\\1\end{pmatrix},$$

因为显然$\mathbf{v}\cdot\mathbf{w}=0$. 该旋转把 $\mathbf{w}$ 映到

$$\mathbf{w}'=R\mathbf{w}=\begin{pmatrix}\frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0\\ \frac{\sqrt{3}}{4} & \frac{3}{4} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}0\\-1\\1\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{3}}{4}\\ \frac{1}{2}\\ -\frac{5}{4}\end{pmatrix}.$$

所以,若 $\mathbf{w}$ 被转过的角为 α,则按 4.2 节的公式(7)有

$$\cos\alpha=\frac{\mathbf{w}\cdot\mathbf{w}'}{|\mathbf{w}||\mathbf{w}'|}=\frac{-\frac{7}{4}}{2}=-\frac{7}{8}. \tag{19}$$

现在,旋转角可以查表得到. 于是,该问题的解答完成了.

下例说明了在涉及长方矩阵的问题中如何应用列向量的记法.

例 10 假定变量 x_1', x_2', x_3' 通过线性法则

$$\left.\begin{aligned}x_1'&=3x_1 \quad\quad + \quad x_3+x_4\\ x_2'&=5x_1+x_2-\quad x_3+x_4\\ x_3'&=\quad x_1 \quad\quad +6x_3+x_4\end{aligned}\right\} \tag{20}$$

与变量 x_1, x_2, x_3, x_4 相联系;变量 x_1'', x_2'' 通过线性法则

$$\left.\begin{aligned}x_1''&=\quad\quad x_1'+2x_2',\\ x_2''&=-3x_1'+\ x_2'+x_3'\end{aligned}\right\} \tag{21}$$

与变量 x_1', x_2', x_3' 相联系. 变量 x_1'', x_2'' 如何与变量 x_1, x_2, x_3, x_4 相联系?

线性法则(20)与(21)的矩阵分别为

$$B=\begin{pmatrix}3 & 0 & 1 & 1\\ 5 & 1 & -1 & 1\\ 1 & 0 & 6 & 1\end{pmatrix} \quad 与 \quad A=\begin{pmatrix}1 & 2 & 0\\ -3 & 1 & 1\end{pmatrix}.$$

我们可以把法则(20)用矩阵乘法写成

$$\mathbf{x}'=B\mathbf{x} \tag{22}$$

的形式，其中

$$\mathbf{x}=\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}, \quad \mathbf{x}'=\begin{pmatrix}x_1'\\ x_2'\\ x_3'\end{pmatrix}.$$

同样，法则(21)也可写成

$$\mathbf{x}''=A\mathbf{x}', \tag{23}$$

其中

$$\mathbf{x}''=\begin{pmatrix}x_1''\\ x_2''\end{pmatrix}.$$

把法则(22)代入(23)，得

$$\mathbf{x}''=A(B\mathbf{x})=(AB)\mathbf{x}. \tag{24}$$

因此，x_1''，x_2'' 也是通过一个线性法则与 x_1, x_2, x_3, x_4 相联系，这个法则的矩阵为

$$AB=\begin{pmatrix}1 & 2 & 0\\ -3 & 1 & 1\end{pmatrix}\begin{pmatrix}3 & 0 & 1 & 1\\ 5 & 1 & -1 & 1\\ 1 & 0 & 6 & 1\end{pmatrix}=\begin{pmatrix}13 & 2 & -1 & 3\\ -3 & 1 & 2 & -1\end{pmatrix}.$$

换言之，

$$x_1''=13x_1+2x_2-x_3+3x_4,$$
$$x_2''=-3x_1+x_2+2x_3-x_4.$$

练　　习

25. 求点 $P(2\sqrt{3}, 2)$ 在到一条直线的正交射影下的像 P'，这条直线

通过原点并与 01 轴成 $\frac{\pi}{3}$ 角。(该射影所对应的矩阵在 9.3 节的例 3(iii)中已给出.)

26. 设 A 是练习 25 提及的射影的矩阵. 试求所有使 $A\mathbf{x}=\mathbf{x}$ 的向量$\mathbf{x}$，并从几何上加以解释. 再求所有使 $A\mathbf{x}=\mathbf{0}$ 的向量 $\mathbf{x}$，这里 $\mathbf{0}$ 是零向量$\begin{pmatrix}0\\0\end{pmatrix}$.

27. 设

$$A=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \mathbf{x}=\begin{pmatrix}x_1\\ x_2\end{pmatrix}.$$

证明，线性方程组 $A\mathbf{x}=\lambda\mathbf{x}$ 当 $\lambda=1$ 或 $\lambda=-1$ 时有非平凡解，而对 λ 的其他值方程组都没有非平凡解.

28. 注意，2.1 节中 n 个未知数 m 个线性方程的方程组(1) 可以简记为 $A\mathbf{x}=\mathbf{b}$ 的形式，其中 A 为 $m\times n$ 系数矩阵，$\mathbf{x}$ 为未知数的列向量($n\times 1$ 矩阵)，而 $\mathbf{b}$ 是常数的列向量($m\times 1$ 矩阵).

29. 用矩阵的方法解 7.4 节练习 29 的旋转问题.

30. 用矩阵的语言表示出下列各变换：

$$\left.\begin{aligned} x_1 &= 3y_1+y_2+y_3\\ x_2 &= -y_1-y_2+y_3\end{aligned}\right\} \quad 与 \quad \left.\begin{aligned} y_1 &= z_1+2z_2\\ y_2 &= -z_1-z_2\\ y_3 &= 2z_1+z_2.\end{aligned}\right\}$$

然后再求用 z_1 与 z_2 来表示 x_1 与 x_2 的线性法则.

第十章 矩阵代数续篇

到现在为止，我们讨论矩阵时一直强调了它们同几何变换的联系. 在后面几章中，我们还要揭示矩阵的许多其他应用，其中有些是与几何毫无关系的. 在我们能够着手叙述和解决这些新的应用之前，还须进一步介绍一些矩阵代数. 我们现在就来做这件事.

10.1 正方矩阵的乘积(续)

恒等矩阵

$n\times n$ 矩阵

$$I=\begin{pmatrix}1 & & & 0\\ & 1 & & \\ & & \ddots & \\ 0 & & & 1\end{pmatrix}$$

(当 $i=j$ 时，它的 ij-元素为 1；当 $i\neq j$ 时，它的 ij-元素为 0) 具有如下性质：对所有 $n\times n$ 矩阵 A，

$$AI=IA=A. \tag{1}$$

它称为 $n\times n$恒等矩阵. 它对应于$\mathscr{R}^n$的恒等变换 $x_i'=x_i(i=1,2,\cdots,n)$. 严格地说，这个 $n\times n$ 矩阵应该用如象 I_n 这样的符号来表示，以区别于 $m\times m$ 恒等矩阵 $I_m(m\neq n)$. 下标 n 通常略去，因为我们所讨论的矩阵究竟多大实际上是很清楚的.

零矩阵

$n\times n$ 矩阵

$$O=\begin{pmatrix}0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix}$$

(它的所有元素都是零)具有如下性质：对所有 $n\times n$ 矩阵 A,

$$AO=OA=O. \tag{2}$$

它称为 $n\times n$ 零矩阵.

矩阵 I 与 O 的性质很像普通算术中的数 1 与 0. 但是，矩阵代数与数的算术之间的相似之处并不是太多的. 举例说吧，矩阵的乘法就不是交换的. 它们之间的另外一些不同之处还可从下例中得以说明.

例 1 找出 2×2 矩阵 A,B,C 与 D,使得

(i) $A^2=O(A\neq O)$;

(ii) $B^2=B(B\neq I$ 或 $O)$;

(iii) $CD=O$(但 $DC\neq O$).

(i) 我们将求 2×2 矩阵的方程 $X^2=O$ 的通解. 设

$$X=\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

则

$$X^2=\begin{pmatrix}a^2+bc & b(a+d)\\ c(a+d) & bc+d^2\end{pmatrix}.$$

方程 $X^2=O$ 等价于四个方程

$$a^2+bc=0,\ b(a+d)=0,\ c(a+d)=0,\ bc+d^2=0.$$

假定 $b\neq 0$,则有 $a+d=0$，于是我们得到解

$$\begin{pmatrix}a & b\\ -\dfrac{a^2}{b} & -a\end{pmatrix}. \tag{3}$$

如果 $b=0$，则 $a=d=0$，这时给出另外一些解

$$\begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix}. \tag{4}$$

矩阵(3)与(4)组成了 $X^2=O$ 的通解. 解的个数是无限的. 把这一点同数的代数中的情况作比较：一个实(或复)系数的二次方程的解不能多于两个. 特别，$x^2=0$ 的解只有 $x=0$.

某些(3)和(4)形式的矩阵有其几何解释. 例如，考虑矩阵

$$A=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5}$$

在通常的直角坐标系中，A 表示到 01 轴的正交射影再接一个绕原点转$\frac{\pi}{2}$的旋转(注意其顺序!). 显然，如果这个合成变换进行两次，所有的向量都映到零. 这就对应了 $A^2=O$ 这一事实.

(ii) $B^2=B$ 的一个特解是

$$B=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{6}$$

它对应到 01 轴的正交射影. 求 $X^2=X$ 的通解留作练习.

(iii) 取上文(5)、(6)两式所定义的矩阵 A 与 B 为例. 此时 $BA=O$ 而 $AB=A$. (其几何解释是什么?)此例亦可用于说明矩阵的乘法一般是不交换的.

练　习

1. 求 2×2 矩阵的方程 $X^2=X$ 的通解.
2. 求 2×2 矩阵的方程 $X^2=I$ 的通解.
3. 从几何上解释例 1 (iii)中的关系式 $BA=O$，$AB=A$.
4. 设 A，B 与 C 都是 2×2 矩阵. 如果 $AB=AC$，是否可以推出 $B=C$?
5. 假定 A，B，C 与 D 都是 2×2 矩阵，并有 $AB=CD$. 是否可以推出对所有 2×2 矩阵 X 有 $AXB=CXD$?

10.2 矩阵的加法与纯量乘法

设 $A=(a_{ij})$, $B=(b_{ij})$ 是两个 $m\times n$ 矩阵．我们定义它们的和为 $m\times n$ 矩阵 (c_{ij})，其中 $c_{ij}=a_{ij}+b_{ij}$．显然，矩阵的加法是交换的：$A+B=B+A$．

除非两个矩阵的行数相同并且列数也相同，不然它们的和是没有定义的．

下面是矩阵加法的两个例子：

$$\begin{pmatrix}2 & 3 & 1\\ 0 & -1 & 1\end{pmatrix}+\begin{pmatrix}1 & -1 & 1\\ 2 & 1 & 2\end{pmatrix}=\begin{pmatrix}3 & 2 & 2\\ 2 & 0 & 3\end{pmatrix};$$

$$\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}+\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}=\begin{pmatrix}x_1+y_1\\ x_2+y_2\\ \vdots\\ x_n+y_n\end{pmatrix}.$$

据此，矩阵相加的规则与向量加法的规则是一致的．

我们还可以进一步说明向量加法与矩阵加法之间的相似性：存在 $m\times n$ 零矩阵 O，使得对所有 $m\times n$ 矩阵 A 都有

$$A+O=O+A=A.$$

矩阵 O 完全由零元素组成．

每一个矩阵 A 都有负矩阵 $-A$，使得

$$A+(-A)=(-A)+A=O.$$

矩阵 $-A$ 是把矩阵 A 的每个元素 a_{ij} 换成 $-a_{ij}$ 而得到的．

矩阵的减法由如下规则定义：

$$A-B=A+(-B).$$

设 k 是一个纯量．我们定义 kA 是把矩阵 A 的每个元素 a_{ij} 换成 ka_{ij} 而得到的矩阵．亦即

$$kA=\begin{pmatrix} ka_{11} & \cdots & ka_{1n} \\ ka_{21} & \cdots & ka_{2n} \\ \vdots & & \vdots \\ ka_{m1} & \cdots & ka_{mn} \end{pmatrix}.$$

特别，$-1\cdot A=-A$.

从上述定义立即可以得到下面的定理.

定理 I 所有 $m\times n$ 实(或复)矩阵的集合在上面所定义的加法与纯量乘法下成为一个向量空间. 这个向量空间的维数是 mn.

证明 显然向量空间的公理都满足. 要证明它的维数是 mn，我们写出其含有 mn 个矩阵的基. 把 ij- 元素为 1 而其余元素都是零的 $m\times n$ 矩阵记为 E_{ij}. 例如，当 $m=2,n=3$ 时，

$$E_{11}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad E_{12}=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad E_{13}=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$E_{21}=\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\quad E_{22}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\quad E_{23}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

由于矩阵的加法与纯量乘法的规则，任一矩阵 $A=(a_{ij})$ 都可以表为这些 E_{ij} 的线性组合. 事实上，

$$A=a_{11}E_{11}+a_{12}E_{12}+\cdots+a_{ij}E_{ij}+\cdots+a_{mn}E_{mn}.$$

再者，这些 E_{ij} 是线性无关的，因为

$$c_{11}E_{11}+\cdots+c_{ij}E_{ij}+\cdots+c_{mn}E_{mn}=O \tag{7}$$

当且仅当对所有的 i,j 都有 $c_{ij}=0$. 要看到这一点，只要注意到(7)式的左边等于矩阵

$$\begin{pmatrix} c_{11} & \cdots & c_{1n} \\ c_{21} & \cdots & c_{2n} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$$

即可. 这就证明了 mn 个矩阵 E_{ij} 组成这个空间的一组基.

例 2 设A是2×2矩阵

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

并且I与O分别是2×2的恒等矩阵与零矩阵．证明，矩阵I，A，A^2，A^3与A^4一定满足形如

$$c_0I+c_1A+c_2A^2+c_3A^3+c_4A^4=O \tag{8}$$

的一个非平凡关系式，其中c_i是不全为零的纯量．

这五个矩阵属于一个四维向量空间（定理 I），所以是线性相关的，因此有形如(8)式的非平凡关系式．证毕．

我们已经证明了非平凡关系式的存在性，但并未具体地把关系式求出来．参看练习 12．

练　　习

6．设

$$A=\begin{pmatrix} 2 & 3 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

写出矩阵$3A$，$-2B$，$3A+2B$，$3A-2B$．

7．如果矩阵A把向量$\mathbf{x}$变换为$\mathbf{x}'$（即$\mathbf{x}'=A\mathbf{x}$），那么矩阵$5A$把$\mathbf{x}$变换为什么向量？

8．证明，矩阵$\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$可以表为一个膨胀与一个旋转的乘积．

9．设A是3×2矩阵，B与C是2×4矩阵．证明，

$$A(B+C)=AB+AC.$$

试推广之．

10．设A与B是2×2矩阵．下式成立与否？请证明你的断言．

$$(A+B)^2=A^2+2AB+B^2.$$

11．设

$$A=\frac{1}{9}\begin{pmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{pmatrix}.$$

证明 $A^2=I$(3×3 恒等矩阵)．可以证明，在直角坐标系中，矩阵 A 表示一个反射．求反射平面的方程．

12．证明，矩阵 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$满足关系式

$$A^2-(a+d)A+(ad-bc)I=O.$$

13．已知 $n\times n$ 矩阵 A 与 B 可交换，证明

$$(A-B)^3=A^3-3A^2B+3AB^2-B^3.$$

当 A 与 B 不能交换时，$(A-B)^3$ 的正确展开式是什么？

10.3 矩阵的转置

$m\times n$ 矩阵 $A=(a_{ij})$ 的转置是一个 $n\times m$ 矩阵，记为 A'，它是把矩阵 A 的行与列互换而得到的．因此

$$A'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

如果我们令 $A'=(a'_{kl})$，则

$$a'_{kl}=a_{lk}\quad(k=1,\cdots,n;l=1,\cdots,m).$$

A' 的 kl-元素等于 A 的 lk-元素．例如，矩阵

$$A=\begin{pmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

的转置是

$$A'=\begin{pmatrix} 1 & -1 \\ 3 & 2 \\ 0 & 1 \end{pmatrix}.$$

注意，对所有矩阵 A 都有 $(A')'=A$．

列向量

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

的转置是行向量 $\mathbf{x}'=(x_1,x_2,\cdots,x_n)$.

对所有 $m\times n$ 矩阵 A，乘积 $A'A$ 与 AA' 都是有定义的. 例如，若 A 如上所示，我们有

$$A'A=\begin{pmatrix}1&-1\\3&2\\0&1\end{pmatrix}\begin{pmatrix}1&3&0\\-1&2&1\end{pmatrix}=\begin{pmatrix}2&1&-1\\1&13&2\\-1&2&1\end{pmatrix}$$

及

$$AA'=\begin{pmatrix}1&3&0\\-1&2&1\end{pmatrix}\begin{pmatrix}1&-1\\3&2\\0&1\end{pmatrix}=\begin{pmatrix}10&5\\5&6\end{pmatrix}.$$

一般地说，若已知 A 是一个 $m\times n$ 矩阵，则 $A'A$ 与 AA' 分别是 $n\times n$ 矩阵与 $m\times m$ 矩阵.

例 3 以转置的语言来描述长度与内积

设 $\mathbf{x}$ 与 $\mathbf{y}$ 是 n 元列向量，则

$$\mathbf{x}'\mathbf{y}=(x_1,\cdots,x_n)\begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix}=x_1y_1+\cdots+x_ny_n. \tag{9}$$

若把 $\mathbf{x}$ 与 $\mathbf{y}$ 看成欧几里得空间 $\mathscr{R}^n$ 中的向量，我们便有 $\mathbf{x}$ 与 $\mathbf{y}$ 的内积的一个表达式

$$\langle\mathbf{x},\mathbf{y}\rangle=\mathbf{x}'\mathbf{y}. \tag{10}$$

特别，

$$\langle\mathbf{x},\mathbf{x}\rangle=\|\mathbf{x}\|^2=\mathbf{x}'\mathbf{x}. \tag{11}$$

在酉空间 $\mathscr{U}^n$(5.6 节)中，这两个公式修改为

$$\langle \mathbf{x}, \mathbf{y} \rangle = \bar{\mathbf{x}}'\mathbf{y} \quad 与 \quad \|\mathbf{x}\|^2 = \bar{\mathbf{x}}'\mathbf{x}, \tag{12}$$

其中 $\bar{\mathbf{x}}$ 是把向量 $\mathbf{x}$ 的每个分量换成它的共轭复数而得到的向量．我们注意到 $\bar{\mathbf{x}}'\mathbf{x}=0$ 当且仅当 $\mathbf{x}$ 是零向量．再者，$\mathbf{x}$ 与 $\mathbf{y}$ 正交当且仅当 $\bar{\mathbf{x}}'\mathbf{y}=0$。

乘积的转置

设 A 与 B 分别为 $m\times n$ 与 $n\times p$ 矩阵，置 $C=AB$。我们要证明 C 的转置由如下公式给出：

$$C'=(AB)'=B'A'. \tag{13}$$

考虑 C' 中的 kl-元素．

$$\begin{aligned} c'_{kl} = c_{lk} &= a_{l1}b_{1k} + a_{l2}b_{2k} + \cdots + a_{ln}b_{nk} \\ &= b'_{k1}a'_{1l} + b'_{k2}a'_{2l} + \cdots + b'_{kn}a'_{nl} \\ &= B'A' \text{ 中的 } kl\text{-元素}. \end{aligned}$$

所以，矩阵 C' 与 $B'A'$ 是一样的．

这一结果是用归纳法证明下述一般公式的第一步：

$$(A_1A_2\cdots A_r)' = A'_r\cdots A'_2A'_1. \tag{14}$$

对称矩阵

一个矩阵 A 如果同自己的转置相等（即 $A=A'$），则称之为对称矩阵．对称矩阵必须是正方的，并且对所有 i,j 都有 $a_{ij}=a_{ji}$。下列前两个矩阵是对称矩阵，但第三个不是：

$$\begin{pmatrix} 2 & 8 \\ 8 & -1 \end{pmatrix}; \begin{pmatrix} a & f & g \\ f & b & h \\ g & h & c \end{pmatrix}; \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

对称矩阵是很重要的，因为它与二次型相联系．我们马上就要讨论这个问题．

练　　习

14．设

$$A=\begin{pmatrix}2 & -1 & 3\\ 1 & 1 & 1\\ 0 & 1 & 2\end{pmatrix},\quad \mathbf{x}=\begin{pmatrix}1\\ 2\\ -1\end{pmatrix}.$$

计算 $A\mathbf{x}$ 与$(A\mathbf{x})'$. 验证$(A\mathbf{x})'=\mathbf{x}'A'$.

15. 举例说明$(AB)'=B'A'$ 及$(ABC)'=C'B'A'$.

16. 证明$(A+B)'=A'+B'$ 与$(kA)'=kA'$.

17. 设

$$A=\begin{pmatrix}1 & 2 & 3\\ 2 & -1 & 1\end{pmatrix}.$$

计算 $A'A$. 注意, $A'A$ 是对称矩阵.

18. 一般性地证明,对任意 $m\times n$ 矩阵 A, $A'A$ 是对称矩阵.

19. 设 $A=(a_{ij})$是实的 $m\times n$ 矩阵, 并假定 $A'A=O$($n\times n$ 零矩阵). 证明,对所有 i,j, $a_{ij}=0$.

20. 证明公式(14).

10.4 二次型与埃尔米特型

所谓 $x_1,\cdots,x_n$ 的二次型是 $x_1,\cdots,x_n$ 的一个多项式,其中每一项都是二次的,亦即,每一项都形如$a_{ij}x_ix_j$,其中 a_{ij} 为常数. 例如

$$3x_1^2+x_3^2-2x_1x_2+x_2x_3$$

就是 x_1,x_2,x_3 的二次型.

二次型出现在数学的许多分支中,例如力学(动能与势能的表达式),几何学(圆锥曲线的方程)以及统计学中(方差与协方差的表达式).

为了看出二次型与矩阵之间的联系,我们来计算 $\mathbf{x}'A\mathbf{x}$,其中

$$\mathbf{x}=\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix},\quad A=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}.$$

我们有

$$Ax=\begin{pmatrix} a_{11}x_1+a_{12}x_2+a_{13}x_3 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3 \\ a_{31}x_1+a_{32}x_2+a_{33}x_3 \end{pmatrix},$$

所以

$$\begin{aligned}\mathbf{x}'A\mathbf{x}=&x_1(a_{11}x_1+a_{12}x_2+a_{13}x_3)+x_2(a_{21}x_1+a_{22}x_2+a_{23}x_3)\\&+x_3(a_{31}x_1+a_{32}x_2+a_{33}x_3).\end{aligned}$$

合并同类项后，我们得到

$$\begin{aligned}\mathbf{x}'A\mathbf{x}=&a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+(a_{12}+a_{21})x_1x_2\\&+(a_{13}+a_{31})x_1x_3+(a_{23}+a_{32})x_2x_3,\end{aligned}$$

我们知道，这是 x_1,x_2,x_3 的二次型.

如果 A 是对称矩阵（对所有 i,j 有 $a_{ij}=a_{ji}$），上面的式子简化为

$$\mathbf{x}'A\mathbf{x}=a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+2a_{13}x_1x_3+2a_{23}x_2x_3. \quad (15)$$

显然，x_1,x_2,x_3 的任意二次型都能表成 $\mathbf{x}'A\mathbf{x}$ 的形式，其中 A 是对称矩阵. 此外，（对称）矩阵 A 的选择是唯一的. 我们称 A 为该二次型的对称矩阵.

上面的讨论可以毫无困难地推广到 n 个变量（n 任意）的二次型. n 个变量的二次型的对称矩阵当然是 $n\times n$ 矩阵.

例 4 求下列二次型的对称矩阵 A：

(i) $2x_1^2+x_3^2-4x_1x_2+6x_2x_3$；

(ii) $ax_1^2+2hx_1x_2+bx_2^2$；

(iii) $(x_1+x_2+x_3)^2$；

(iv) $3x_2^2-x_1x_3-2x_2x_4$.

(i) $a_{11}=2$，$a_{22}=0$，$a_{33}=1$，$2a_{12}=-4$，$2a_{13}=0$，$2a_{23}=6$.

因此，

$$A=\begin{pmatrix} 2 & -2 & 0 \\ -2 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

(ii) $A=\begin{pmatrix} a & h \\ h & b \end{pmatrix}$.

(iii) $(x_1+x_2+x_3)^2=x_1^2+x_2^2+x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$，所以

$$A=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

(iv) $$A=\begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 3 & 0 & -1 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

$x_1,\cdots,x_n$ 的二次型 $\mathbf{x}'A\mathbf{x}$ 可以用 $\mathscr{R}^n$ 的内积的语言方便地表达出来.设 $\mathbf{x}$ 是 $\mathscr{R}^n$ 的向量(写成列向量),而 A 是实的 $n\times n$ 矩阵，从 10.3 节的例 3 马上得知

$$\mathbf{x}'A\mathbf{x}=\langle\mathbf{x},A\mathbf{x}\rangle.$$

二次型是如下型式的特殊情况：

$$\mathbf{x}'A\mathbf{y}=\langle\mathbf{x},A\mathbf{y}\rangle=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j, \tag{16}$$

这个型式称为 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 的双线性型. 它可以看成是 $x_1,\cdots,x_n$ 的线性表达式，也可以看成是 $y_1,\cdots,y_n$ 的线性表达式,这一点解释了它的名称.

我们下文所需要的一个重要关系式是

$$\langle\mathbf{x},A\mathbf{y}\rangle=\langle A'\mathbf{x},\mathbf{y}\rangle. \tag{17}$$

证明 $\langle A'\mathbf{x},\mathbf{y}\rangle=(A'\mathbf{x})'\mathbf{y}=\mathbf{x}'(A')'\mathbf{y}=\mathbf{x}'A\mathbf{y}=\langle\mathbf{x},A\mathbf{y}\rangle$.

关系式(17)对任意的实 $n\times n$ 矩阵都成立. 如果 A 是对称矩阵,则

$$\langle \mathbf{x}, A\mathbf{y}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle.$$

上文的讨论可以作一个重要的推广，即考虑用 $\mathscr{U}^n$ 的内积表示的复双线性型. 下面我们就讨论这种推广. 设 A 是 $n\times n$ 复矩阵，而$\langle\ ,\ \rangle$表示 $\mathscr{U}^n$ 的内积，则

$$\langle \mathbf{x}, A\mathbf{y}\rangle = \overline{\mathbf{x}}' A\mathbf{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x}_i y_j. \tag{18}$$

这是 $\overline{x}_1,\cdots,\overline{x}_n$ 与 $y_1,\cdots,y_n$ 的复双线性型. 下面的关系式成立：

$$\langle \mathbf{x}, A\mathbf{y}\rangle = \langle \overline{A}'\mathbf{x}, \mathbf{y}\rangle, \tag{19}$$

这里 $\overline{A}$ 表示 A 的共轭，即把每个 a_{ij} 换成 $\overline{a}_{ij}$ 而得到的. 这个关系式可以用类似于证明关系式 (17) 的方法来证明. 注意，关系式(17)是(19)的特殊情况，即 A 是实矩阵 (故 $\overline{A}=A$) 且 $\mathbf{x}$ 与 $\mathbf{y}$ 都是实向量的情况.

一个 $n\times n$ 复矩阵如果满足 $\overline{A}'=A$ （A 等于其共轭转置)，则称为埃尔米特矩阵. 一个实矩阵是埃尔米特矩阵当且仅当它是对称矩阵. 如果 A 是埃尔米特矩阵，则

$$\langle \mathbf{x}, A\mathbf{y}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle;$$

特别，

$$\langle \mathbf{x}, A\mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{x}\rangle.$$

当 A 是埃尔米特矩阵时，表达式$\langle \mathbf{x}, A\mathbf{x}\rangle$称为埃尔米特型. 这就推广了实二次型的概念.

例 5 证明，埃尔米特型仅取实值

$$\overline{\langle \mathbf{x}, A\mathbf{x}\rangle} = \langle A\mathbf{x}, \mathbf{x}\rangle \text{(根据内积的共轭对称性)}$$

$$= \langle \mathbf{x}, A\mathbf{x}\rangle \text{(因为 } A \text{ 是埃尔米特矩阵)}.$$

由此可见$\langle \mathbf{x}, A\mathbf{x}\rangle$等于自己的共轭，所以是实数.

例 6 下列矩阵中哪些是埃尔米特矩阵？

(i) $\begin{pmatrix} 1 & i \\ i & -2 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$; (iii) $\begin{pmatrix} i & 3 \\ 3 & 1 \end{pmatrix}$;

$$\text{(iv)} \begin{pmatrix} 5 & 4-3i & -2-6i \\ 4+3i & 5 & 2-6i \\ -2+6i & 2+6i & 8 \end{pmatrix}.$$

$A=(a_{ij})$是埃尔米特矩阵当且仅当$\bar{A}'=A$，亦即，对所有的i,j有$\bar{a}_{ji}=a_{ij}$. 特别，$\bar{a}_{ii}=a_{ii}$，所以埃尔米特矩阵的所有对角元素都必须是实的. 由此可见，矩阵(iii)不是埃尔米特矩阵. 矩阵(i)也不是，因为$\bar{a}_{12}\neq a_{21}$. 剩下的两个矩阵是埃尔米特矩阵.

例 7(说明例5的实例) 写出埃尔米特型$\langle \mathbf{x}, A\mathbf{x}\rangle$，其中

$$A=\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}.$$

$$A\mathbf{x}=\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} x_1+ix_2 \\ -ix_1+2x_2 \end{pmatrix},$$

于是，

$$\langle \mathbf{x}, A\mathbf{x}\rangle=\bar{\mathbf{x}}'A\mathbf{x}=(\bar{x}_1, \bar{x}_2)\begin{pmatrix} x_1+ix_2 \\ -ix_1+2x_2 \end{pmatrix}$$

$$=\bar{x}_1x_1+(i\bar{x}_1x_2-ix_1\bar{x}_2)+2\bar{x}_2x_2.$$

这是实的，因为$\bar{x}_1x_1$与$\bar{x}_2x_2$是实的，而括号里的项是$z+\bar{z}$的形式

后面的若干章节中我们解某些动力学问题时，要用到本节的一些结果. 在另外一些领域中，我们还可以发现这些结果的其他应用. 例如，推广到某些酉空间的埃尔米特型在量子理论中起重要的作用. 下面的参考书曾讨论到这一课题：R. A. Newing 与 J. Cunningham, *Quantum Mechanics*, Oliver & Boyd, 1968.

练　　习

21．求下列二次型的对称矩阵：

(i) $x_1^2-2x_1x_2+x_2^2$;

(ii) $4x_2^2+x_1x_3$;

(iii) $x_1^2+2x_3^2-4x_1x_3+x_2x_3$.

22. 对练习 21 的每一个对称矩阵，写出双线性型$\langle \mathbf{x}, A\mathbf{y}\rangle$.

23. 设 $A=\begin{pmatrix} 1 & 2 \\ -2 & 2 \end{pmatrix}$. 计算$\langle \mathbf{x}, A\mathbf{y}\rangle$,$\langle A\mathbf{x}, \mathbf{y}\rangle$与$\langle A'\mathbf{x}, \mathbf{y}\rangle$.

24. 设 A 与 B 是 $n\times n$ 对称矩阵. 下列各矩阵中哪些一定是对称矩阵?

(i) $A+B$; (ii) $A-B$; (iii) $3A$; (iv) A^2; (v) AB.

25. 一个矩阵称为斜称矩阵当且仅当 $A'=-A$. 下列各矩阵中哪些是斜称矩阵:

$$\text{(i)}\begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & -1 \\ -3 & 1 & 0 \end{pmatrix};\quad \text{(ii)}\begin{pmatrix} 1 & 2 & 3 \\ -2 & 0 & -1 \\ -3 & 1 & 0 \end{pmatrix}.$$

证明，在斜称矩阵中，所有对角元素 a_{ii} 都是零.

26. 设 A 是任意 $n\times n$ 矩阵. 证明 $A+A'$ 是对称矩阵而 $A-A'$ 是斜称矩阵. 举数字的例子说明之.

27. 证明，任一正方矩阵 A 能够表成 $A=B+C$，使得 B 是对称矩阵而 C 是斜称矩阵.

[提示：用练习 26.]

28. C 是斜称矩阵. 证明，对所有的 $\mathbf{x}$ 都有 $\mathbf{x}'C\mathbf{x}=0$.

29. 下列矩阵中哪些是埃尔米特矩阵?

$$\text{(i)}\ A=\begin{pmatrix} 0 & 1+i \\ 1-i & 2 \end{pmatrix};\quad \text{(ii)}\ B=\begin{pmatrix} i & 1 \\ 1 & 2 \end{pmatrix};$$

$$\text{(iii)}\ C=\begin{pmatrix} 1 & i & -2i \\ -i & 0 & 2+i \\ 2i & 2-i & 2 \end{pmatrix}.$$

30. 设 A 与 B 如练习29，再设

$$\mathbf{x}=\begin{pmatrix} i \\ 1+i \end{pmatrix},\quad \mathbf{y}=\begin{pmatrix} 1 \\ 2i \end{pmatrix}.$$

在 $\mathscr{U}^2$ 中计算:

(i) $\langle \mathbf{x}, A\mathbf{y}\rangle$; (ii) $\langle A\mathbf{x}, \mathbf{y}\rangle$; (iii) $\langle \mathbf{x}, B\mathbf{y}\rangle$; (iv) $\langle B\mathbf{x}, \mathbf{y}\rangle$; (v) $\langle \bar{B}'\mathbf{x}, \mathbf{y}\rangle$. 再证明$\langle \mathbf{x}, A\mathbf{x}\rangle$是实数，但$\langle \mathbf{x}, B\mathbf{x}\rangle$不是实数.

第十一章 矩阵的逆

11.1 引论

假定A与B是两个$n\times n$矩阵,使

$$AB=BA=I, \tag{1}$$

这里I是$n\times n$恒等矩阵. 我们说B是A的逆矩阵,记为$B=A^{-1}$. 请把这个定义同逆变换的定义作比较.

如果A有一个逆矩阵B,则B是它唯一的逆矩阵. 证明:假如C也是满足

$$AC=CA=I$$

的矩阵,则得到

$$C=IC=(BA)C=B(AC)=BI=B.$$

一个正方矩阵不一定有逆矩阵. $n\times n$零矩阵O就没有逆矩阵,因为不可能找到一个矩阵B使$OB=I$. 显然,表示线性变换T的矩阵有逆矩阵当且仅当T有逆变换. 因此,9.3节例3中的矩阵(i),(ii),(iv),(v)与(vi)有逆矩阵,但矩阵(iii)没有.

一个正方矩阵如果有逆矩阵,则称为非奇异的;如果没有逆矩阵,则称为奇异的.

非正方的矩阵的逆矩阵是没有定义的.

本章中我们要研究各种方法,以判断一个已知的正方矩阵是否为奇异矩阵. 我们将发现,行初等变换(最初在2.2节介绍)的方法对此目的是十分有用的. 下一节中,我们要进一步考察这种变换.

例1 证明,若A是非奇异矩阵,则A^{-1}也是非奇异矩阵.

置$B=A^{-1}$. 从关系式(1)我们看到不但B是A的逆矩阵,而

且 A 也是 B 的逆矩阵，亦即 $(A^{-1})^{-1}=A$。

例 2 假设 C 是非奇异矩阵，而矩阵 A 使 $CA=I$。证明，A 是非奇异的。

我们有 $C^{-1}(CA)=C^{-1}I$，因此 $A=C^{-1}$。根据例 1，A 是非奇异的。

练 习

1．用几何观察的方法写出 9.3 节例 3 中各非奇异矩阵的逆矩阵。对每一情况，再把原矩阵和它的逆矩阵相乘，以验证你的答案.

2．假定 $AB=AC$ 且 A 是非奇异的。证明 $B=C$。

3．已知 A 与 B 是 $n\times n$ 非奇异矩阵，证明，AB 也是非奇异矩阵，并写出它的逆矩阵.

[提示：参考 6.2 节.]

4．求 k 个非奇异矩阵的积 $A_1A_2\cdots A_k$ 的逆矩阵.

5．设 A 与 B 是 $n\times n$ 非奇异矩阵。下列矩阵中哪些一定是非奇异的？

(i) $-A$；(ii) $5A$；(iii) $A+B$；(iv) A^2。

6．求 $\begin{pmatrix}3 & -4\\ 4 & 3\end{pmatrix}$ 的逆矩阵.

7．假定 B 是非奇异矩阵。如果 $BX=ABC$，能否推出 $X=AC$？求出 X 的正确表达式.

8．验证下面八个矩阵在矩阵乘法下成为一个群.

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\ \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix},\ \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix},\ \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},$$

$$\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix},\ \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\ \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\ \begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix}.$$

[提示：把这些矩阵看成为旋转与反射，再回忆我们对对称群所进行的讨论.]

9．以下面的坐标系为参照(图 78)，把等边三角形的对称群表示成为矩阵的群.

10．在量子力学中，泡利自旋矩阵

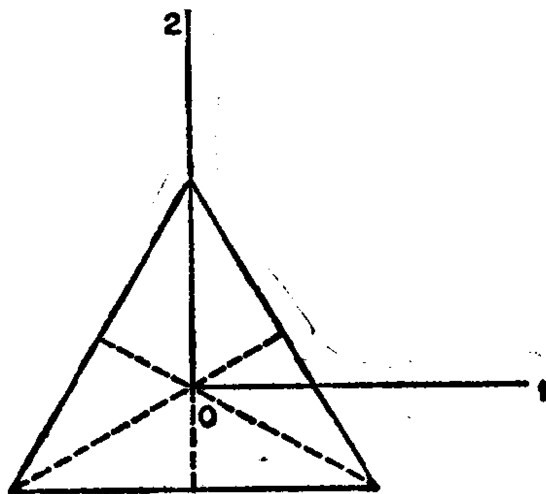


图 78

$$\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

用于描述电子的自旋．证明 $\sigma_x^2=\sigma_y^2=\sigma_z^2=I$，然后写出这些矩阵的逆矩阵．再证明 $\sigma_x\sigma_y=-\sigma_y\sigma_x=i\sigma_z$，并求出另外两个类似的关系式．

证明，σ_x 与 σ_z 在矩阵乘法下生成一个 8 阶的群．由 σ_x，σ_y 与 σ_z 生成的群的阶是多少？

11．证明，矩阵

$$A=\begin{pmatrix}-1&0\\0&1\end{pmatrix}\quad 与\quad B=\begin{pmatrix}-1&1\\0&1\end{pmatrix}$$

满足 $A^2=B^2=I$．证明，它们在矩阵乘法下生成一个无限阶的群．

[提示：考虑 BA 的幂．]

12．证明，矩阵

$$A=\begin{pmatrix}5&2&-4\\2&8&2\\-4&2&5\end{pmatrix}$$

满足矩阵方程 $A^2-9A=O$，这里 O 是零矩阵．由此再证明 A 是奇异矩阵．

11.2 初等矩阵

设 A 是 $m\times n$ 矩阵．在 2.2 节中，我们介绍了三类可施于 A 的行初等变换．为了易于参考，我们把这些变换重写在这儿：

类型 I：交换两行；

类型 II：把一行乘以一个非零常数；

类型 III：把一行的 k 倍加于另一行.

我们现在作一重要的注记：要把任意一个行初等变换施行于 A，只要在 A 的左边乘上适当的 $m\times m$ 矩阵 E.

例 3

我们以 A 是 4×3 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}$$

的情况来说明上面的注记.

类型 I：要把 A 的第二行与第四行交换，只要把 A 左乘以矩阵

$$E_1=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

因为

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

类型 II：要把 A 的第三行乘以 -2，只要把 A 左乘以矩阵

$$E_2=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

因为

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}
= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -2a_{31} & -2a_{32} & -2a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}.
$$

类型 III：要把 A 的第四行的 k 倍加于第一行，只要把 A 左乘以矩阵

$$
E_3 = \begin{pmatrix} 1 & 0 & 0 & k \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
$$

因为

$$
\begin{pmatrix} 1 & 0 & 0 & k \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}
= \begin{pmatrix} a_{11}+ka_{41} & a_{12}+ka_{42} & a_{13}+ka_{43} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}.
$$

显然，矩阵 A 有多少列是无关紧要的——上面所给的左乘因子 E_i 在所有 $4\times n$ 矩阵上实现行变换，而不论 n 是什么数值。

像上面的 E_1, E_2 与 E_3 这样的矩阵，只要用它左乘，便可实现行初等变换．这类矩阵称为初等矩阵．根据它们所实现的行初等变换的类型，我们又分别称它们为类型 I，类型 II 或类型 III 的初

等矩阵.

一个初等矩阵必须是一个正方矩阵．如果它是 $m\times m$ 矩阵，它当然只能左乘在 $m\times n$ 矩阵上．而 $m\times n$ 矩阵 A 上的每一个行初等变换，都对应唯一的初等矩阵 E，它以左乘（亦即 EA）实现这个变换.

有一个简单的办法求出 E 来．考察 $E=EI$，这里 I 是 $m\times m$ 恒等矩阵．由此可以看出，E 是由对恒等矩阵施行所给的行初等变换而得到的．请以上面所讨论的三个例子来验证这个规则．下面再举一例.

例 4　求对应于下列行初等变换的 5×5 初等矩阵：

(i) 把第二行与第五行交换；

(ii) 把第二行的 k 倍加于第五行.

设想把这两个变换施于 5×5 恒等矩阵

$$\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{pmatrix},$$

我们便得到

$$\text{(i)}\quad E_1=\begin{pmatrix}1&0&0&0&0\\0&0&0&0&1\\0&0&1&0&0\\0&0&0&1&0\\0&1&0&0&0\end{pmatrix};$$

$$\text{(ii)}\quad E_2=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&k&0&0&1\end{pmatrix}.$$

请验证施行初等变换(i)与(ii)于任意 $5\times n$ 矩阵 A 后，得到的矩阵确实分别是 E_1A 与 E_2A。

现在我们给出初等矩阵的一个重要性质。

定理 I 所有初等矩阵都是非奇异的，并且它们的逆矩阵也是初等矩阵。

证明 这一结论可从行初等变换都是可逆的这一事实得出。假定 E 与 E^* 分别是对应于某个行初等变换及其逆变换的初等矩阵。此时，对所有 $m\times n$ 矩阵 A 都有 $E^*EA=A$。此外，亦不难看出，$EE^*A=A$。令 $A=I$，我们得到

$$E^*E=EE^*=I,$$

所以 $E^*=E^{-1}$。

例 5 求例 3 中各初等矩阵的逆矩阵

$E_1^{-1}=E_1$。一般地说，对应于类型 I 的初等变换的矩阵总是等于自己的逆矩阵。

要消除 E_2(对 A)的作用，我们必须把(E_2A 的)第三行乘以 $-\frac{1}{2}$。因此，

$$E_2^{-1}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -\frac{1}{2} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

最后，E_3 的逆矩阵所对应的变换必须是从第一行减去第四行的 k 倍。因而，

$$E_3^{-1}=\begin{pmatrix}1 & 0 & 0 & -k\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

定理 II 有限个初等矩阵的乘积是非奇异的。

证明 参看 11.1 节的练习 4。

在本节结束之前，我们还要用初等矩阵的语言复述一下行等价的定义，并用它来复述早先的一个结果。

定理 III $A \underset{\sim}{r} B$ 当且仅当存在一组初等矩阵 $E_1, \cdots, E_k$，使得 $B = E_k \cdots E_1 A$。

定理 IV 一个 $m \times n$ 矩阵 A 总可以表为

$$A = E_1 \cdots E_k R$$

的形式，其中 $E_1, \cdots, E_k$ 是 $m \times m$ 初等矩阵，而 R 是一个 $m \times n$ 简化阶梯形矩阵，并且 $A \underset{\sim}{r} R$。

证明 根据 2.3 节及上面的定理 III，我们能够找到一个简化阶梯形矩阵 R 和一组初等矩阵 $F_1, \cdots, F_k$，使 $R = F_k \cdots F_1 A$，所以 $A = F_1^{-1} \cdots F_k^{-1} R$。又因为初等矩阵的逆矩阵也是初等矩阵。定理证毕。

练 习

13. 写出对应下列 $3 \times n$ 矩阵上的行初等变换的初等矩阵：

(i) 把第三行乘以 -2；

(ii) 把第二行的 -3 倍加于第一行；

(iii) 交换第一行与第三行。

14. 写出练习 13 各矩阵的逆矩阵，并说明它们所对应的行初等变换。

15. 把矩阵

$$A = \begin{pmatrix} 1 & 2 & -3 & -8 \\ 1 & 1 & 2 & 0 \end{pmatrix}$$

表成 $A = E_1 \cdots E_k R$ 的形式，其中 $E_1, \cdots, E_k$ 为初等矩阵，而 R 是简化阶梯形矩阵。

16. 证明，给定任意一个 $m \times n$ 矩阵 A，则存在一个非奇异矩阵 C，使 CA 是简化阶梯形矩阵。

17. 设 $A = (a_{ij})$ 是 $m \times n$ 矩阵。把它的行记成行向量 $\mathbf{a}_1 = (a_{11} a_{12} \cdots$

a_{1n}), $\mathbf{a}_2=(a_{21}a_{22}\cdots a_{2n})$，等等．我们可以用这些向量把$A$写成

$$A=\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\\ \vdots\\ \mathbf{a}_m\end{pmatrix}.$$

证明，如果$B=(b_{ij})$是$r\times m$矩阵，则乘积BA是$r\times n$矩阵

$$BA=\begin{pmatrix}b_{11}\mathbf{a}_1+b_{12}\mathbf{a}_2+\cdots+b_{1m}\mathbf{a}_m\\ b_{21}\mathbf{a}_1+b_{22}\mathbf{a}_2+\cdots+b_{2m}\mathbf{a}_m\\ \cdots\quad\cdots\quad\cdots\quad\cdots\\ b_{r1}\mathbf{a}_1+b_{r2}\mathbf{a}_2+\cdots+b_{rm}\mathbf{a}_m\end{pmatrix}.$$

因此，BA的行是A的行的线性组合．

用正文中出现的矩阵按此方法构作矩阵的乘积BA，特别，就B是初等矩阵的情况作此练习．

11.3 在正方矩阵中的应用

现在假定A是正方的$m\times m$矩阵，并且$A\underset{\sim}{r}R$，其中R是简化阶梯形矩阵．则只有两种情况：

(i) R的每一列都含有行的首元素．因为R是正方矩阵，仅有的可能是$R=I$($m\times m$恒等矩阵)；或者

(ii) R中某一列不含有行的首元素．在这种情况下，R中行的首元素一定不是逐级地延伸到右边，而是跳过某一列．所以R的最后一行仅含零元素．下面的5×5简化阶梯形矩阵说明了这种情况：

$$\begin{pmatrix}1&0&2&0&0\\0&1&3&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{pmatrix}.$$

我们要证明，A是非奇异的当且仅当上述第一种可能性出现，

亦即 $A \underset{\sim}{r} I$。我们需要一个引理.

引理 I 设R是一个矩阵，它的最后一行仅含零元素. 再设M是一个矩阵，使乘积 RM 有定义. 则 RM 的最后一行也只含零元素.

证明 用 11.2 节的练习 17.

定理 V 一个正方矩阵A是非奇异的当且仅当A行等价于恒等矩阵.

证明 先假定 $A \underset{\sim}{r} I$. 根据上一节的定理 IV，此时存在一组初等矩阵 $E_1, \cdots, E_k$，使 $A = E_1 \cdots E_k$. 所以，$A^{-1} = E_k^{-1} \cdots E_1^{-1}$，亦即$A$是非奇异的.

反过来，假定A不是行等价于I. 此时 $A \underset{\sim}{r} R$，这里R是个简化阶梯形矩阵，且它的最后一行仅含零元素. 于是，存在一组初等矩阵 $E_1, \cdots, E_k$，使

$$A = E_1 \cdots E_k R, \tag{2}$$

所以

$$R = E_k^{-1} \cdots E_1^{-1} A. \tag{3}$$

根据假定，A有逆矩阵 A^{-1}. 把(3)式的两边顺次右乘以 $A^{-1}, E_1, \cdots, E_k$，我们得到

$$R(A^{-1} E_1 \cdots E_k) = I. \tag{4}$$

据引理，(4)式左边是一个最后一行仅含零元素的矩阵，因此不等于恒等矩阵. 这与(4)式矛盾. 证毕.

我们可以把上述定理用初等矩阵的语言叙述如下.

定理 V 的推论 一个正方矩阵A是非奇异的当且仅当它可表为初等矩阵的乘积：

$$A = E_1 \cdots E_k.$$

此外，如果A是奇异的，则它可表为 $A = E_1 \cdots E_k R$ 的形式，这里 $E_1, \cdots, E_k$ 是初等矩阵，而R是一个简化阶梯形矩阵，并且其最后一行仅含零元素.

定理 V 提供了矩阵非奇异性的一个简单判别法：用行初等变换把矩阵 A 化为简化阶梯形矩阵 R. 如果 $R=I$，则 A 是非奇异的；若 $R\neq I$，则 A 是奇异的.

如果 A 有逆矩阵，则它可在上述判别过程中作为副产品而算出来. 假定 A 是非奇异的，而 $F_1,\cdots,F_k$ 是初等矩阵，依次施行它们所对应的行变换，就把 A 变成 I. 换句话说，

$$I=F_k\cdots F_1A.$$

所以

$$A^{-1}=F_k\cdots F_1=F_k\cdots F_1I. \tag{5}$$

由此可见，(依次)施行把 A 化为 I 的那些行变换，就把 I 化为 A^{-1}.

例 6 求

$$A=\begin{pmatrix}1 & 1 & -3\\ -1 & 0 & 2\\ -3 & 5 & 0\end{pmatrix}$$

的逆矩阵.

我们希望用行变换同时变化 A 与 I，因此，把这两个矩阵写在一道是很方便的：

$$\left(\begin{array}{ccc|ccc} \multicolumn{3}{c}{A} & \multicolumn{3}{c}{I}\\ 1 & 1 & -3 & 1 & 0 & 0\\ -1 & 0 & 2 & 0 & 1 & 0\\ -3 & 5 & 0 & 0 & 0 & 1\end{array}\right).$$

现在我们把这个迭置的矩阵化为简化阶梯形(如表 10 所示).

把 A 化为 I 的同一过程也把 I 化为 A^{-1}. 因此，

$$A^{-1}=\begin{pmatrix}10 & 15 & -2\\ 6 & 9 & -1\\ 5 & 8 & -1\end{pmatrix}.$$

表 10

变化过程	矩阵
以第一行的首元素为主元，清理第一列（用两个类型III 的变换）	$\begin{pmatrix}1&1&-3&1&0&0\\0&1&-1&1&1&0\\0&8&-9&3&0&1\end{pmatrix}$
以第二行的首元素为主元，清理第二列（用两个类型III 的变换	$\begin{pmatrix}1&0&-2&0&-1&0\\0&1&-1&1&1&0\\0&0&-1&-5&-8&1\end{pmatrix}$
标准化第三行	$\begin{pmatrix}1&0&-2&0&-1&0\\0&1&-1&1&1&0\\0&0&1&5&8&-1\end{pmatrix}$
以第三行的首元素为主元，清理第三列	$\begin{pmatrix}1&0&0&10&15&-2\\0&1&0&6&9&-1\\0&0&1&5&8&-1\end{pmatrix}$

例 7 矩阵

$$A=\begin{pmatrix}2&-3&1\\2&3&-2\\4&0&-1\end{pmatrix}$$

是非奇异的吗？

用行初等变换，我们求得

$$Ax\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}.$$

因而 A 是奇异的.

下面的定理在下一章中有一定的重要性.

定理 VI 设 A 是 $n\times n$ 矩阵，$\mathbf{x}$ 是列向量

$$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}.$$

则由 n 个关于 $x_1,\cdots,x_n$ 的线性方程组成的齐次方程组

$$A\mathbf{x}=\mathbf{0} \tag{6}$$

有非平凡解当且仅当 A 是奇异的.

证明 先假定 A 是奇异的. 此时, 存在一个简化阶梯形矩阵 R, 它的最后一行仅由零元素组成, 并使 $A \underset{r}{\sim} R$. 根据 2.3 节的定理 I, 方程组

$$R\mathbf{x}=\mathbf{0} \tag{7}$$

与方程组(6)同解. 假定 R 有 r ($<n$)个非零行. 根据 2.4 节的定理 III, 方程组(7)的通解中有 $n-r(>0)$ 个自由变量. 因此, 方程组(7)有非平凡解, 从而方程组(6)也有非平凡解.

反过来, 如果 A 是非奇异的, 则 $A\mathbf{x}=\mathbf{0}$ 蕴涵着

$$\mathbf{x}=A^{-1}(A\mathbf{x})=A^{-1}\mathbf{0}=\mathbf{0}.$$

因此, 在这种情况下方程组(6)没有非平凡解. 证毕.

最后, 我们叙述一个定理, 它把 $\mathscr{R}^n$ 中线性无关性的判别法(3.2 节)与本节的讨论联系起来.

定理 VII 设 A 是 $n\times n$ 实矩阵. A 是非奇异的当且仅当 A 的行(看作 $\mathscr{R}^n$ 的向量)线性无关.

对于复矩阵, 类似的定理也成立.

为判定一个矩阵是否有逆矩阵, 本节已提出了若干充要条件. 此外还有一个条件*, 并且可能是最有用的条件. 它要牵涉到行列式的概念, 我们接下去就要介绍这一概念了.

练　习

18. 下列矩阵中哪些是非奇异的? 如果是非奇异的, 求它的逆矩阵.

(i) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$; (ii) $\begin{pmatrix} 3 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & 1 & -1 \end{pmatrix}$; (iii) $\begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 4 & 1 & 3 \\ -1 & 1 & 1 & 0 \\ -2 & 2 & -5 & -1 \end{pmatrix}$.

---

* 参看 12.2 节的定理 VI.

19. 把练习18中的非奇异矩阵表为初等矩阵的乘积。

20. 求矩阵

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

的逆矩阵。一个正方矩阵如果每行每列都正好只有一个元素为1，其余元素全是零，则称之为置换矩阵。你能说明用此名称的理由吗？

第十二章　行　列　式

12.1　引论与定义

行列式最初是在 n 个未知数 n 个线性方程的理论中引进的．考虑 $n=2$ 的情况．只要 $a_{11}a_{22}-a_{12}a_{21}\neq 0$，我们用消去法得知，方程组

$$\begin{aligned} a_{11}x_1+a_{12}x_2&=b_1, \\ a_{21}x_1+a_{22}x_2&=b_2 \end{aligned} \tag{1}$$

有唯一的解

$$x_1=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}},$$

$$x_2=\frac{b_2a_{11}-b_1a_{12}}{a_{11}a_{22}-a_{12}a_{21}}.$$

表达式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21} \tag{2}$$

称为系数矩阵 $A=(a_{ij})$ 的行列式．它通常记成(2)式那样，或者记为 $|A|$ 或 $\det A$．一个 2×2 矩阵的行列式称为二阶行列式．

例 1　设 $A=(a_{ij})$ 是 2×2 矩阵，B 是交换 A 的两列而得到的矩阵．证明 $|B|=-|A|$．

$$B=\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix}=\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

所以 $|B|=b_{11}b_{22}-b_{12}b_{21}=a_{12}a_{21}-a_{11}a_{22}=-|A|$．下面的命题也是正确的：如果交换 A 的两行，所得到的矩阵的行列式等于 $-|A|$．

例 2　证明，齐次方程组

$$\begin{aligned} a_{11}x_1+a_{12}x_2=0,\\ a_{21}x_1+a_{22}x_2=0 \end{aligned} \tag{3}$$

有非平凡解当且仅当$|A|=0$.

设$|A|=a_{11}a_{22}-a_{12}a_{21}=0$. 我们先处理$A=O$(零矩阵)的情况. 此时$x_1$与$x_2$的任意值都是方程组(3)的解. 当$A\neq O$时,$A$中某个元素(例如$a_{12}$)非零. 我们希望方程组有如下的解:

$$x_1=k,\quad x_2=-\frac{a_{11}k}{a_{12}}\ (k\text{ 任意}).$$

这组值显然是(3)的第一个方程的解. 它也是第二个方程的解,因为从$|A|=0$推得

$$\begin{aligned} &a_{21}k+a_{22}\left(-\frac{a_{11}k}{a_{12}}\right)\\ &\qquad=(a_{12}a_{21}-a_{11}a_{22})\frac{k}{a_{12}}=-|A|\frac{k}{a_{12}}=0. \end{aligned}$$

如果$a_{12}=0$而A中其他某个元素非零,也可同样论证.

反过来,若$|A|\neq 0$,我们能够得到这个方程组的唯一(平凡的)解,正如本节开头所指出的那样.

再考虑$n=3$的情况. 我们问:齐次方程组

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3=0,\\ a_{21}x_1+a_{22}x_2+a_{23}x_3=0,\\ a_{31}x_1+a_{32}x_2+a_{33}x_3=0 \end{aligned}\right\} \tag{4}$$

满足什么条件时才有非平凡解?假定方程组(4)有非平凡解

$$x_1=k_1,\quad x_2=k_2,\quad x_3=k_3,$$

设其中$k_3\neq 0$. 把x_1,x_2与x_3的这组值代入方程组(4),再用后两个方程把k_1与k_2用k_3表出. 我们得到

$$(a_{21}a_{32}-a_{31}a_{22})k_1+(a_{23}a_{32}-a_{22}a_{33})k_3=0$$

与

$$(a_{22}a_{31}-a_{32}a_{21})k_2+(a_{23}a_{31}-a_{21}a_{33})k_3=0.$$

这两个方程可用二阶行列式写成

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} k_1 = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} k_3,$$
$$\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} k_2 = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} k_3. \tag{5}$$

把方程组 (4) 的第一个方程乘以$\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$,并把(5)的两个方程代入,得到

$$\left\{ a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \right\} k_3 = 0. \tag{6}$$

因为我们假定 $k_3 \neq 0$,所以得知花括号里的式子必须等于零. 这个式子就称为 3×3 矩阵 $A=(a_{ij})$的行列式:

$$|A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{7}$$

把(7)式展开,我们得到

$$|A| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \tag{8}$$

3×3 矩阵的行列式称为三阶行列式.

我们曾假定 $k_3 \neq 0$. 当 $k_3 = 0$ 而如果 $k_2 \neq 0$ 时,也没有必要把这些计算从头再做一遍,而只要把上述的列指标 2 与 3 互换即可. 我们得到方程(6)的稍加修改的形式,由此找出使方程组(4) 有非平凡解且 $x_2 = k_2 \neq 0$ 的条件. 这个条件是

$$a_{11}\begin{vmatrix} a_{23} & a_{22} \\ a_{33} & a_{32} \end{vmatrix} - a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} + a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = 0. \tag{9}$$

把此式与(7)式对照,我们看到,(9) 式的左边等于$-|A|$. 同样,

假定 $k_1 \neq 0$ 也得出 $-|A|=0$ 的结论．因此，我们证明了，如果方程组(4)有非平凡解，则 $|A|=0$．其逆命题也是正确的(参看 13.1 节的定理 I)．

我们现在希望把所引进的二阶与三阶行列式的定义加以推广，而给出 n 阶行列式的定义．带着这个问题，让我们再来仔细考察由(8)式定义的三阶行列式．我们作如下注记：

(i) 每一项都是 $a_{1i_1}a_{2i_2}a_{3i_3}$ 的形式，其中列指标 i_1,i_2,i_3 是 1,2,3 的一个排列．如果一项的行指标是按自然顺序 1,2,3 出现(象上面所写的那样)，我们就说这一项写成了标准形式．

(ii) 一共有六项．当每一项都写成标准形式时，这些项代表了列指标所有可能的排列．

(iii) A 的主对角线元素的乘积的项 $a_{11}a_{22}a_{33}$ 带正号．

(iv) 任意两项如果差别仅在于交换了两个列指标，则它们所带的符号相反．例如，$a_{11}a_{23}a_{32}$ 是由 $a_{11}a_{22}a_{33}$ 交换列指标 2 与 5 而得到的，所以带负号．如果我们再把 $a_{11}a_{23}a_{32}$ 的列指标 1 与 2 交换，便得到 $a_{12}a_{23}a_{31}$．这一项带正号．

(v) 一般地说，$a_{1i_1}a_{2i_2}a_{3i_3}$ 带正号还是负号，取决于它的列指标的排列 i_1,i_2,i_3 是由自然顺序 1, 2, 3 经过偶数次还是奇数次交换而得到的．

这些注记可以用对列指标 1,2,3 的置换来叙述．

(i) 每一项都是 $a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}$ 的形式，这里 π 是对 1,2,3 的一个置换，称为与 $a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}$ 相关的置换．明确地说，与 $a_{1i_1}a_{2i_2}a_{3i_3}$ 相关的置换是

$$\pi=\begin{pmatrix}1 & 2 & 3\\ i_1 & i_2 & i_3\end{pmatrix}.$$

例如，与 $a_{13}a_{21}a_{32}$ 相关的置换是

$$\pi=\begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix},$$

这个置换可以用循环的记号写成 $\pi=(132)$.

(ii) 一共 $3!$ 项,代表了对符号 1,2,3 的所有可能的置换.

(iii) 与恒等置换相关的项带正号.

(iv) 如果 τ 是一个对换(交换两个符号),则与 π 相关的项和与 $\tau\pi$ 相关的项带相反的符号. 这两项的差别是交换了两个列指标. 明确地说,如果 τ 交换了符号 i 与 j,则 $a_{1\tau\pi(1)}a_{2\tau\pi(2)}\cdot a_{3\tau\pi(3)}$是由$a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}$交换列指标 i 与 j 而得到的. 例如,$a_{1\underline{2}}\ a_{2\underline{3}}a_{31}$是与$\pi=(123)$相关的. 交换列指标 2 与 3,我们得到 $a_{1\underline{3}}\ a_{2\underline{2}}a_{31}$(交换了加线的指标). 这一项是与 $\tau\pi$ 相关的,这里$\tau=(23)$. 验证:$\tau\pi=(23)(123)=(13)$.

(v) 与 π 相关的项带正号还是负号,取决于 π 是偶置换还是奇置换.

下表中我们列出了三阶行列式$|A|$中所出现的项.

表 11

项	相关的置换	奇偶性	项的符号
$a_{11}a_{22}a_{33}$	恒等置换	偶	+
$a_{11}a_{23}a_{32}$	(23)	奇	−
$a_{12}a_{21}a_{33}$	(12)	奇	−
$a_{12}a_{23}a_{31}$	(123)	偶	+
$a_{13}a_{21}a_{32}$	(132)	偶	+
$a_{13}a_{22}a_{31}$	(13)	奇	−

我们现在能够给 n 阶行列式(也就是 $n\times n$ 矩阵的行列式)下定义了.

定义 I 设 $A=(a_{ij})$是 $n\times n$ 矩阵. A 的行列式定义为

$$|A|=\sum_{\pi}\pm a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}, \tag{10}$$

这里的和是对 $1,2,\cdots,n$ 上的所有 $n!$ 个置换 π 取的,每项前面带

正号还是负号，取决于π是偶置换还是奇置换.

例 3 在5×5矩阵$A=(a_{ij})$的行列式$|A|$的展开式中，出现下面哪些项？带什么符号？

$$\text{(i)}\ a_{11}a_{23}a_{31}a_{44}a_{52};\qquad \text{(ii)}\ a_{13}a_{24}a_{31}a_{45}a_{52};$$

$$\text{(iii)}\ a_{25}a_{41}a_{33}a_{54}a_{12}.$$

(i)项不出现，因为符号1作为列指标出现两次. (ii)项出现. 其相关的置换为

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 3 & 4 & 1 & 5 & 2\end{pmatrix}=(13)(245)=(13)(25)(24).$$

这是奇置换，所以(ii)项带负号.

(iii)项也出现. 它的标准形式为$a_{12}a_{25}a_{33}a_{41}a_{54}$. 相关的置换为

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 5 & 3 & 1 & 4\end{pmatrix}=(1254)=(14)(15)(12).$$

这是奇置换，所以(iii)项也是带负号的.

例 4 求行列式

$$\begin{vmatrix}a & b & c & d\\ b & a & c & c\\ b & c & a & d\\ b & c & d & d\end{vmatrix}$$

展开式中b^2cd项的系数.

把上面行列式的i行j列交叉点上的元素记为f_{ij}. 因为b^2是b^2cd的因子，我们只要考虑涉及$f_{12}f_{21}$，$f_{12}f_{31}$或$f_{12}f_{41}$的项. 其中第一种可能性可以排除，因为在此行列式中没有形如$f_{12}f_{21}f_{3-}f_{4-}$的项可以等于b^2cd. 据此，我们得出此行列式展开式中同类于b^2cd的项，如表12所示. 可见，在展开式中b^2cd的系数是$+1-1-1=-1$.

表 12

同类于 b^2cd 的项	相关的置换（用循环记号写出）	奇偶性	符号
$f_{12}f_{23}f_{31}f_{44}$	(123)	偶	+
$f_{12}f_{24}f_{31}f_{43}$	(1243)	奇	−
$f_{12}f_{23}f_{34}f_{41}$	(1234)	奇	−

练　习

1．在 6×6 矩阵 $A=(a_{ij})$ 的行列式 $|A|$ 的展开式中，出现下面哪些项？带什么符号？

(i) $a_{14}a_{23}a_{32}a_{46}a_{55}a_{61}$；

(ii) $a_{34}a_{63}a_{25}a_{16}a_{42}a_{51}$；

(iii) $a_{13}a_{25}a_{31}a_{43}a_{56}a_{62}$。

2．求值：

(i) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$；　(ii) $\begin{vmatrix} 3 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & 1 & -1 \end{vmatrix}$；　(iii) $\begin{vmatrix} a & b & c \\ 0 & 0 & 0 \\ d & e & f \end{vmatrix}$；

(iv) $\begin{vmatrix} a & 0 & d \\ b & 0 & e \\ c & 0 & f \end{vmatrix}$；　(v) $\begin{vmatrix} 2 & 3 & -1 \\ 0 & 1 & 6 \\ 0 & 0 & 5 \end{vmatrix}$；

(vi) $\begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{vmatrix}$.

3．证明，方程组

$$x_1+2x_2+3x_3=0,$$
$$2x_1+3x_2+4x_3=0,$$
$$3x_1+4x_2+5x_3=0$$

有非平凡解。

4. 证明，如果 A 是 $n\times n$ 矩阵，k 是纯量，则

$$|kA|=k^n|A|.$$

12.2 行列式的行变换

用 n 阶行列式 $|A|$ 的定义来求它的值是一个冗长的过程，它牵涉到 $n!$ 项的计算．本节中我们将发现，用行初等变换的工具把 A 化简为阶梯形式后，$|A|$ 的值就很容易求出来了．这一方法还有一个长处：它揭示了行列式与线性方程及矩阵的一般理论之间的联系．

我们先考虑某些较易求其行列式值的矩阵．

定理 I 设 A 是 $n\times n$ 矩阵且它有一行全是零，则 $|A|=0$．

证明 假定 A 的第 i 行全是零，则对任意置换 π，$a_{i\pi(i)}=0$．所以，$|A|$ 的展开式中每一项都含有因子零．

例 5 求 $|A|$ 的值，其中

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix}.$$

考虑 $|A|$ 展开式中的非零项 $a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}a_{4\pi(4)}$．因为 $a_{41}=a_{42}=a_{43}=0$，所以一定要 $\pi(4)=4$．由于 π 是 1,2,3,4 的置换，得知 $\pi(3)\neq 4$，所以一定要 $\pi(3)=3$．如此继续下去，我们看到，$|A|$ 中唯一的非零项是 $a_{11}a_{22}a_{33}a_{44}$．这一项带正号，因此，$|A|=a_{11}a_{22}a_{33}a_{44}$．

上面这个矩阵 A 是上三角矩阵的一个例子．上三角矩阵就是主对角线下方的所有元素都是零的矩阵．类似地，一个矩阵如果其主对角线上方的所有元素都是零，则称为下三角矩阵．上三角矩阵或下三角矩阵都可简称为三角矩阵．注意，如果一个矩阵既是上三角矩阵，又是下三角矩阵，则它只能在主对角线上有非零元

素。这样的矩阵称为对角矩阵.

不难把例 5 推广成下面的定理.

定理 II $n\times n$ 三角矩阵 $A=(a_{ij})$ 的行列式值等于其主对角线元素的乘积，亦即

$$|A|=a_{11}a_{22}\cdots a_{nn}.$$

特别, $|I|=1$, 这里 I 表示恒等矩阵.

我们已经知道任意一个正方矩阵都能够用行初等变换化简为阶梯形矩阵. 根据定理 II, 阶梯形矩阵的行列式值容易求出(阶梯形矩阵都是三角矩阵). 我们现在要研究行初等变换对行列式值的影响，然后得出把一个行列式化简为阶梯形式来求出它的值的方法. 为此目的，考虑如下一些例子.

例 6 比较下面两个矩阵的行列式：

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ y_1 & y_2 & y_3 & y_4 & y_5 \end{pmatrix} \quad 与 \quad B=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ x_1 & x_2 & x_3 & x_4 & x_5 \end{pmatrix}.$$

矩阵 B 是把矩阵 A 的第三行与第五行交换而得到的，我们的记法强调了这一事实. 注意，$|A|$ 的展开式中的任意一项在 $|B|$ 的展开式中也出现，反之亦然. 考虑 $|A|$ 中具体的一项，例如 $t=a_{12}a_{24}x_3a_{45}y_1$, 我们把它写成了标准形式($x$ 与 y 所带的下标是列指标). 在 $|B|$ 中, 同一项的标准形式是 $a_{12}a_{24}y_1a_{45}x_3$. 因此，在 $|A|$ 中 t 与置换 $\pi=(1245)$ 相关, 而在 $|B|$ 中它与置换 $\tau\pi$ 相关, 其中 τ 是对换(13). 可见 t 在 $|A|$ 与 $|B|$ 的展开式出现时带相反的符号.

这一论证过程可以用于证明，$|A|$ 展开式中的每一项在 $|B|$ 展开式中出现时带相反的符号. 所以 $|B|=-|A|$.

容易把这一例子推广成如下的重要结果：如果 B 是把正方矩

阵A的两行交换而得到的矩阵，则$|B|=-|A|$。

例 7　求行列式的值：

$$|A|=\begin{vmatrix} 1 & 3 & 2 & 5 \\ 2 & 1 & 3 & 4 \\ 8 & 7 & -6 & 3 \\ 2 & 1 & 3 & 4 \end{vmatrix}.$$

不需要展开！只要注意到$|A|$的第二行与第四行相同，把它们交换并不改变任何东西．但是，根据上面的结果，这一交换产生的行列式的值是$-|A|$．因此，$|A|=-|A|$，故$|A|=0$。

一般地说，若A是任意一个有两行相同的矩阵，则$|A|=0$。

例 8　证明

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ ca_{21} & ca_{22} & ca_{23} & ca_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = c\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

由定义，左边的行列式等于

$$\sum_{\pi} \pm a_{1\pi(1)}(ca_{2\pi(2)})a_{3\pi(3)}a_{4\pi(4)}$$

$$=c\sum_{\pi} \pm a_{1\pi(1)}a_{2\pi(2)}a_{3\pi(3)}a_{4\pi(4)},$$

结论得证．

例 9　证明

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ p_1+q_1 & p_2+q_2 & p_3+q_3 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ p_1 & p_2 & p_3 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ q_1 & q_2 & q_3 \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

p与q所带的下标是列指标．左边行列式的展开式是

$$\sum_{\pi} \pm a_{1\pi(1)}(p_{\pi(2)}+q_{\pi(2)})a_{3\pi(3)}$$

$$=\sum_{\pi} \pm a_{1\pi(2)}p_{\pi(2)}a_{3\pi(3)} + \sum_{\pi} \pm a_{1\pi(1)}q_{\pi(2)}a_{3\pi(3)}$$

$=$右边两个行列式的和.

例 10 设 $A=(a_{ij})$ 是 $n\times n$ 矩阵,证明

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1}+ka_{i1} & a_{j2}+ka_{i2} & \cdots & a_{jn}+ka_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=|A|.$$

用与例 9 类似的论证过程,我们得知,左边的行列式等于

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$=|A|+k$(一个 i 行与 j 行相等的行列式)

$=|A|$(根据例 7 后面的说明).

刚才完成的这些例题说明了行初等变换对行列式的影响. 我们把所得到的结果总结成下面的定理.

定理 III 假定把下面的行初等变换施于 $n\times n$ 矩阵 A 上:

类型 I: 交换 A 的两行,得到 A_1;

类型 II: 把 A 的一行乘以 c,得到 A_2;

类型 III: 把 A 的一行的 k 倍加于另一行,得到 A_3.

则

$$|A_1|=-|A|,\quad |A_2|=c|A|,\quad |A_3|=|A|.$$

定理 II 与定理 III 为求行列式的值提供了简捷的方法. 为了

求$|A|$，我们先用行初等变换把A化为阶梯形式. 根据定理 III，每一步所得的行列式是与$|A|$密切相关的. 最后的阶梯矩阵是个三角矩阵，它的行列式值可以用定理 II 求出.

例 11 求$|A|$的值，其中

$$A=\begin{pmatrix}0 & 2 & 3\\ 2 & 4 & -2\\ 3 & 4 & 5\end{pmatrix}.$$

反复应用定理 III，我们得到如表 13 所示的计算过程. 最后一个行列式是三角的.

所以，根据定理 II，

$$|A|=-2\cdot1\cdot2\cdot11=-44.$$

表 13

行列式	变换方法
$|A|=-\begin{vmatrix}2 & 4 & -2\\ 0 & 2 & 3\\ 3 & 4 & 5\end{vmatrix}$	把 1、2 行交换.
$=-2\begin{vmatrix}1 & 2 & -1\\ 0 & 2 & 3\\ 3 & 4 & 5\end{vmatrix}$	提出因子 2.
$=-2\begin{vmatrix}1 & 2 & -1\\ 0 & 2 & 3\\ 0 & -2 & 8\end{vmatrix}$	把 1 行的 -3 倍加于 3 行.
$=-2\begin{vmatrix}1 & 2 & -1\\ 0 & 2 & 3\\ 0 & 0 & 11\end{vmatrix}$	把 2 行加于 3 行.

例 12 求$|A|$的值，其中

$$A=\begin{pmatrix}2 & -3 & 1\\ 2 & 3 & -2\\ 4 & 0 & -1\end{pmatrix}.$$

仿上一个例题，我们得到

$$|A|=\begin{vmatrix} 2 & -3 & 1 \\ 0 & 6 & -3 \\ 0 & 6 & -3 \end{vmatrix}=0,$$

后一等号成立是根据例 7 后面的说明.

例 12 的矩阵 A 最初是在 11.3 节的例 7 中出现的. 那里证明了 A 是奇异矩阵. 比较这两个例题，你便可想像出奇异矩阵与它的行列式之间的一般关系了. 下文的定理 VI 就叙述这种关系. 现在先让我们分析一下初等矩阵与行列式的关系.

定理 IV 当 E 是类型 I,类型 II(把一行乘以 $c\neq 0$)或类型III 的初等矩阵时, E 的行列式值分别为 -1, c 或 1.

证明 把矩阵 A 的两行交换，相当于左乘一个类型 I 的初等矩阵 E, 根据定理 III, $|EA|=-|A|$. 特别,置 $A=I$,便得到

$$|E|=|EI|=-|I|=-1.$$

定理的其余两种情况同理可证. 注意,在所有情况下, $|E|\neq 0$.

定理 V 设 E 是一个初等矩阵,则 $|EA|=|E||A|$.

证明 设 E 是类型 I 的. 此时,

$$\begin{aligned} |EA| &= -|A| \quad &(\text{据定理 III}) \\ &= |E||A| \quad &(\text{据定理 IV}). \end{aligned}$$

其余两种情况同理可证.

定理 V 的推论 如果 $E_1, E_2, \cdots, E_n$ 为初等矩阵,则

$$|E_1E_2\cdots E_nA|=|E_1||E_2|\cdots|E_n||A|. \tag{11}$$

证明 据定理 V,

$$|E_1(E_2\cdots E_nA)|=|E_1||E_2\cdots E_nA|.$$

再重复应用定理 V,结论即得.

定理 VI 一个正方矩阵 A 是非奇异的当且仅当 $|A|\neq 0$.

证明 由 11.2 节的定理 IV, A 可以表成

$$A=E_1\cdots E_nR \tag{12}$$

的形式，其中 $E_1,\cdots,E_n$ 为初等矩阵而 R 是简化阶梯形矩阵.

先假定 A 是非奇异的. 此时，根据 11.3 节的定理 V，A 表成 (12) 的形式时 $R=I$. 所以

$$|A|=|E_1|\cdots|E_n||I|.$$

这是一组非零数（据定理 IV）的积，所以 $|A|\neq 0$.

反过来，如果 A 是奇异的，它表成 (12) 的形式时 R 的最后一行仅含零元素，因此 $|R|=0$（据定理 I）. 由此可见 $|A|=0$.

乘积的行列式

上文中的定理是许多行列式理论问题的基础. 我们要用这些定理来证明一个重要结果：两个矩阵乘积的行列式等于行列式的积.

定理 VII $|AB|=|A||B|$.

证明 把 A 表成 (12) 的形式：$A=E_1\cdots E_nR$. 根据 (11) 式，

$$|AB|=|E_1\cdots E_nRB|=|E_1|\cdots|E_n||RB|. \tag{13}$$

如果 A 是非奇异的，我们有 $R=I$. 此时 $A=E_1\cdots E_n$，于是

$$|A|=|E_1|\cdots|E_n|.$$

又，$RB=IB=B$. 因此，(13) 式变成

$$|AB|=|A||B|.$$

如果 A 是奇异的，则 R 的最后一行仅含零元素. 由 11.3 节的引理 I，RB 的最后一行也仅含零元素. 因此，$|RB|=0$. 从而由 (13) 式得到 $|AB|=0$. 又因为 $|A|=0$（据定理 VI），我们也有 $|A||B|=0$. 证毕.

定理 VII 的推论 $|A_1A_2\cdots A_r|=|A_1||A_2|\cdots|A_r|$.

证明 反复应用定理 VII.

例 13 设

$$A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix},\quad B=\begin{pmatrix} c & -d \\ d & c \end{pmatrix}.$$

求 $|A|$，$|B|$ 与 $|AB|$ 的值，并导出 a,b,c,d 的一个恒等式.

$$AB=\begin{pmatrix}ac-bd & -(ad+bc)\\ ad+bc & ac-bd\end{pmatrix},$$

而$|A||B|=|AB|$. 展开以后我们得到

$$(a^2+b^2)(c^2+d^2)=(ac-bd)^2+(ad+bc)^2.$$

这个恒等式说明了，如果两个整数都能表成整数的平方和，则它们的积也能表成整数的平方和。例如，$13=2^2+3^2, 29=2^2+5^2$，则 $13\cdot29=377=11^2+16^2$.

练　　习

5. 从定理 II 出发，直接证明定理 IV.

6. 用行初等变换求下列行列式的值：

(i) $\begin{vmatrix}1 & 6 & 7\\ 3 & 4 & -9\\ -2 & -5 & 8\end{vmatrix}$;　　(ii) $\begin{vmatrix}1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\\ 1 & 1 & 1 & 1\end{vmatrix}$;

(iii) $\begin{vmatrix}2 & 5 & -3 & -2\\ -2 & -3 & 2 & -5\\ 1 & 3 & -2 & 2\\ -1 & -6 & 4 & 3\end{vmatrix}$.

7. 试求，当 λ 为何值时，方程组

$$-2x_1+x_2=\lambda x_1,$$
$$x_1-2x_2=\lambda x_2$$

有非平凡解. 对每一个这样的 λ，求方程组的通解.

[提示：把所有含 x_1, x_2 的项归到左边，再用 12.1 节的例 2.]

8. 证明，一组正方矩阵的乘积 $A_1A_2\cdots A_r$ 是非奇异的当且仅当每个 A_i ($i=1,\cdots,r$)都是非奇异的.

9. 求下述行列式的值，把结果表成因子的积.

$$\begin{vmatrix}1 & a & a^2\\ 1 & b & b^2\\ 1 & c & c^2\end{vmatrix}.$$

10. A 与 B 都是正方矩阵，且 $AB=I$(恒等矩阵). 证明 A 与 B 可交

12.3 行列式的列变换

我们现在要证明，一个矩阵的行列式等于它的转置的行列式。这个结果很重要，因为它使我们也可以变换行列式的列来求它的值，就像变换它的行来求值一样。

我们先考虑初等矩阵的转置的行列式。

定理 VIII 设 E 是任意初等矩阵，E'为它的转置，则$|E'|=|E|$.

证明 类型 I 与类型 II 的矩阵 E 都等于自己的转置 E'. 类型 III 的初等矩阵是主对角线上都是 1 的三角矩阵，E' 具有同样的性质，所以二者的行列式值都是1.由此可见，在所有情况下都有$|E'|=|E|$.

定理 IX 设正方矩阵 A 中有一列元素全是零，则$|A|=0$.

证明 在$|A|$的展开式中，每一项都含有来自零元素列的因子零.

定理 X 对任一正方矩阵 A, $|A'|=|A|$.

证明 根据 11.3 节定理V的推论，A可以表为$A=E_1\cdots E_nR$的形式，当 A 非奇异时，$R=I$；当 A奇异时，R 有一行全是零. 由12.2节定理 VII 的推论，

$$|A|=|E_1|\cdots|E_n||R|. \tag{14}$$

再者，$A'=R'E'_n\cdots E'_1$(据 10.4 节的公式(14))，所以

$$\begin{aligned}|A'|&=|R'||E'_n|\cdots|E'_1|\\&=|R'||E_n|\cdots|E_1|\end{aligned} \tag{15}$$

(据定理 VIII). 又，当 $R=I$ 或 R 有一行全是零时，都有$|R'|=|R|$: 在前一情况下这一等式是显然的，因为 $I'=I$. 在后一情况下，R'有一列全是零，所以 R 与 R' 的行列式值都是零.

把(14)式与(15)式相比，我们看到，在任何情况下都有$|A'|=$

$|A|$.

定理 X 的推论 把 12.2 节定理 III 中的“行”全部换成“列”，定理仍然成立.

例 14

$$\begin{vmatrix} 4 & 1 & -3 \\ -12 & 0 & 9 \\ 8 & 3 & -6 \end{vmatrix} = \begin{vmatrix} 4 & 1 & -3+\frac{3}{4}\cdot 4 \\ -12 & 0 & 9+\frac{3}{4}(-12) \\ 8 & 3 & -6+\frac{3}{4}\cdot 8 \end{vmatrix}$$

$$= \begin{vmatrix} 4 & 1 & 0 \\ -12 & 0 & 0 \\ 8 & 3 & 0 \end{vmatrix} = 0.$$

在此，只用了一个类型 III 的列变换就足以把行列式的值求出来了.

练　　习

11．求行列式的值：

$$\begin{vmatrix} 3 & 2 & 3 \\ 6 & 0 & 6 \\ -3 & -1 & -3 \end{vmatrix}.$$

12．设 A 是斜称矩阵

$$A=\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}.$$

先证明 $|A|=-|A'|$，再推出 $|A|=0$.

13．一般性地证明，奇数阶的斜称矩阵的行列式值是零.

12.4　把行列式按一行(一列)展开

再考虑 $n\times n$ 矩阵 $A=(a_{ij})$ 的行列式的定义 I (12.1节). 每

一项正好含有第一行的一个元素为因子，亦即，正好含有 $a_{11}, a_{12}, \cdots, a_{1n}$ 这些元素中的一个为因子．含 a_{11} 的项正好是使 $\pi(1)=1$ 的那些项．它们的总和是

$$a_{11}\sum_{\pi} \pm a_{2\pi(2)}\cdots a_{n\pi(n)}, \tag{16}$$

这里的和取遍 $1,2,\cdots,n$ 上的 $(n-1)!$ 个使 $\pi(1)=1$ 的置换 π，而每一项带正号还是负号，取决于 π 是偶置换还是奇置换．(16)式中出现的 π 事实上是 $2,3,\cdots,n$ 上的置换．据此，我们看出(16)式中的和 $\sum\limits_{\pi}$ 等于 $(n-1)$ 阶行列式

$$M_{11}=\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix},$$

它是把 $|A|$ 的第一行与第一列取掉而得到的．

一般地说，把 A 的第 i 行与第 j 列取掉去得到的 $|A|$ 的子行列式记为 M_{ij}，它称为 $|A|$ 的余子式．

$|A|$ 的展开式是这样开头的：

$$|A|=a_{11}M_{11}+\cdots.$$

现在考虑 $|A|$ 中含 a_{12} 的项．先交换第一列与第二列，把 a_{12} 换到1,1位置．

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}=-\begin{vmatrix} a_{12} & a_{11} & a_{13} & \cdots & a_{1n} \\ a_{22} & a_{21} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n2} & a_{n1} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

$$=-a_{12}\begin{vmatrix} a_{21} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \cdots & a_{nn} \end{vmatrix}+\text{与 } a_{12} \text{ 无关的项}$$

$$=-a_{12}M_{12}+\text{与 } a_{12} \text{ 无关的项}.$$

为了求$|A|$中含a_{13}的项的总和,先把a_{13}换到第二列(交换第二列与第三列),再把它换到第一列(交换第一列与第二列),从而使它到达位置(1,1).用此方法我们保证了原第三列外的各列仍按原顺序排列.每次列的交换各产生一个负号,因此

$$|A|=+\begin{vmatrix} a_{13} & a_{11} & a_{12} & a_{14} & \cdots \\ a_{23} & a_{21} & a_{22} & a_{24} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ a_{n3} & a_{n1} & a_{n2} & a_{n4} & \cdots \end{vmatrix}=+a_{13}M_{13}+\text{与 } a_{13} \text{ 无关的项}.$$

如此继续下去,我们得到

$$|A|=a_{11}M_{11}-a_{12}M_{12}+a_{13}M_{13}-\cdots+(-1)^{1+n}a_{1n}M_{1n}. \quad (17)$$

展开式(17)称为$|A|$按第一行的展开式.这就推广了12.1节的(7)式.

类似地,我们可以得到按第一列的展开式

$$|A|=a_{11}M_{11}-a_{21}M_{21}+a_{31}M_{31}-\cdots+(-1)^{n+1}a_{n1}M_{n1}. \quad (18)$$

更一般地说,我们能够得到$|A|$按任意一行或任意一列的展开式.这些展开式最好用$|A|$的代数余子式A_{ij}写出,A_{ij}只是余子式带上适当的符号:

$$A_{ij}=(-1)^{i+j}M_{ij}. \quad (19)$$

从M_{ij}到A_{ij},所添加的符号可以在如下阵列的第i行与第j列的交叉点上找到:

$$
\begin{pmatrix}
+ & - & + & - & \cdots \\
- & + & - & + & \cdots \\
+ & - & + & - & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{pmatrix}. \tag{20}
$$

不难证明，

$$
|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}a_{in} \tag{21}
$$

（按第 i 行的展开式）

与

$$
|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \tag{22}
$$

（按第 j 列的展开式）

例 15 按第二列展开，求行列式

$$
|A| = \begin{vmatrix}
1 & 2 & 3 & -1 \\
5 & 0 & -1 & 3 \\
-2 & 0 & 0 & 2 \\
0 & -3 & 2 & 0
\end{vmatrix}
$$

的值.

从展开式(22)及(19)式(或符号矩阵(20))，我们得到

$$
|A| = -a_{12}M_{12} + a_{22}M_{22} - a_{32}M_{32} + a_{42}M_{42}
$$

$$
= -2\begin{vmatrix} 5 & -1 & 3 \\ -2 & 0 & 2 \\ 0 & 2 & 0 \end{vmatrix} + 0 + 0 + (-3)\begin{vmatrix} 1 & 3 & -1 \\ 5 & -1 & 3 \\ -2 & 0 & 2 \end{vmatrix}
$$

$$
= -2(-2)\begin{vmatrix} 5 & 3 \\ -2 & 2 \end{vmatrix} + (-3)\left\{-2\begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} + 2\begin{vmatrix} 1 & 3 \\ 5 & -1 \end{vmatrix}\right\}
$$

（两个三阶行列式都按第三行展开）

$$
= 4(10+6) + 6(9-1) - 6(-1-15) = 208.
$$

求例 15 的行列式的值的方法有多种. 上面就是一个很好用的方法. 下面的方法也可供选择.

$$\begin{vmatrix} 1 & 2 & 3 & -1 \\ 5 & 0 & -1 & 3 \\ -2 & 0 & 0 & 2 \\ 0 & -3 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & -10 & -16 & 8 \\ 0 & 4 & 6 & 0 \\ 0 & -3 & 2 & 0 \end{vmatrix}$$ (以类型 III 的行变换清理第一列)

$$= \begin{vmatrix} -10 & -16 & 8 \\ 4 & 6 & 0 \\ -3 & 2 & 0 \end{vmatrix}$$ (按第一列展开)

$$= 8 \begin{vmatrix} 4 & 6 \\ -3 & 2 \end{vmatrix}$$ (按第三列展开)

$$= 8 \cdot 26 = 208.$$

例 16 证明

$$D = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)(c-a).$$

把D展开并分解因子，容易验证这个公式．但是，并非一定要这样做．我们把D看成 a,b,c 的多项式，可以证实它必须含有因子$(a-b)$，因为若在D中置 $a=b$，则行列式的前两行完全相同，它的值变成零．因此，根据余数定理，D被$(a-b)$整除．同理，D有因子$(b-c)$与$(c-a)$。所以，

$$D = k(a-b)(b-c)(c-a).$$

k 一定是与 a，b，c 无关的常数，因为在D的展开式中没有次数高于 3 的项出现．最后，D中 bc^2 的系数等于 1，因为 bc^2 除作为主对角线元素的乘积出现外，就不再出现了．由此可见 $k=1$.

例 17 证明，n 阶行列式

$$|A_n|=\begin{vmatrix} 2\cos\theta & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 2\cos\theta & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 2\cos\theta & 1 & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & 1 & 2\cos\theta \end{vmatrix}$$

满足递推关系式

$$|A_{n+2}|-2\cos\theta|A_{n+1}|+|A_n|=0, \tag{23}$$

据此再证明

$$|A_n|=\begin{cases} \dfrac{\sin(n+1)\theta}{\sin\theta} & (\text{当 } \sin\theta\neq 0 \text{ 时}) \\ n+1 & (\text{当 } \theta=2k\pi \text{ 时,其中 } k \text{ 为整数}) \\ (-1)^n(n+1) & (\text{当 } \theta=(2k+1)\pi \text{ 时}). \end{cases}$$

(在动力学的一个问题中，这是一个重要的行列式. 参看 D. E.Rutherford, *Classical Mechanics*, Oliver & Boyd, Third Edition, 1964, §93.)

把$|A_{n+2}|$($n+2$ 阶行列式,其中 $n+2\geqslant 3$)按第一行展开.

$$|A_{n+2}|=2\cos\theta|A_{n+1}|-1\begin{vmatrix} 1 & 1 & 0 & & \cdots & 0 \\ 0 & 2\cos\theta & 1 & & & 0 \\ 0 & 1 & 2\cos\theta & & & \cdot \\ 0 & 0 & 1 & & & \cdot \\ \vdots & \vdots & \vdots & & & \cdot \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\theta \end{vmatrix}$$

$$=2\cos\theta|A_{n+1}|-|A_n|,$$

关系式(23)得证. 我们利用关系式(23)对 n 归纳证明$|A_n|$的公式. 首先,公式对 $n=1$ 与 $n=2$ 都成立,其证明留作练习. 请注意,在能够用关系式(23)求$|A_3|$之前,$|A_1|$与$|A_2|$都必须知道,所以,在这个归纳法证明中,只验证 $n=1$ 的情况是不够的.

现在假定我们已经证明了当 $\sin\theta\neq 0$ 时,

$$|A_m|=\frac{\sin(m+1)\theta}{\sin\theta} \quad 与 \quad |A_{m+1}|=\frac{\sin(m+2)\theta}{\sin\theta}.$$

此时，根据关系式(23)，

$$|A_{m+2}|=2\cos\theta\frac{\sin(m+2)\theta}{\sin\theta}-\frac{\sin(m+1)\theta}{\sin\theta}.$$

再者，

$$2\sin A\ \cos B=\sin(A+B)+\sin(A-B),$$

所以

$$2\sin(m+2)\theta\ \cos\theta=\sin(m+3)\theta+\sin(m+1)\theta.$$

由此得出

$$|A_{m+2}|=\frac{\sin(m+3)\theta}{\sin\theta}.$$

这样，我们用归纳法完成了 $\sin\theta\neq 0$ 的情况的证明．其他情况也可以用归纳法证明；或者把它们看成当 θ 趋近于 π 的倍数时的极限，这样，过程会更简短一些．

练　　习

14．用不同的方法求本章所出现的行列式的值．

15．求行列式的值：

$$\text{(i)}\begin{vmatrix} 1 & 6 & 4 & 7 \\ 3 & 4 & 2 & -9 \\ 0 & 0 & 2 & 0 \\ -2 & -5 & -1 & 8 \end{vmatrix}; \qquad \text{(ii)}\begin{vmatrix} 1 & 0 & 6 & 7 \\ 6 & 3 & 3 & 2 \\ 3 & 0 & 4 & -9 \\ -2 & 0 & -5 & 8 \end{vmatrix}.$$

[提示：按适当的行或列展开．]

16．证明，$n+1$ 阶行列式

$$D_{n+1}(x)=\begin{vmatrix} x & a & a & \cdots & a \\ a & x & a & \cdots & a \\ \cdots & \cdots & \cdots & \cdots & \\ a & a & a & \cdots & x \end{vmatrix}$$

含有因子 $(x-a)^n$．求其另一个因子，并写出 $D_{n+1}(x)$ 的因子分解式．

17．把例 17 的解答补充完整．

18．试求，当 x 取何值时，行列式

$$\begin{vmatrix} x & b & c & d \\ b & x & c & d \\ b & c & x & d \\ b & c & d & x \end{vmatrix}$$

等于零？

19．设 A_{ij} 为 n 阶行列式 $|A|$ 的代数余子式．

(i) 证明

$$x_1A_{k1}+x_2A_{k2}+\cdots+x_nA_{kn}=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ x_1 & x_2 & \cdots & x_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

即把 $|A|$ 的第 k 行换成 $(x_1,x_2,\cdots,x_n)$ 后得到的行列式．

(ii) 由此推出，若 $i\neq k$，则

$$a_{i1}A_{k1}+a_{i2}A_{k2}+\cdots+a_{in}A_{kn}=0.$$

20．一个 $n\times n$ 矩阵 A 的伴随矩阵定义为代数余子式的矩阵

$$\mathrm{adj}A=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

证明，$A\mathrm{adj}A=|A|I$，其中 I 为 $n\times n$ 恒等矩阵．

[提示：用练习 19．]

由此推出，非奇异矩阵 A 的逆矩阵可由下式给出：

$$A^{-1}=\frac{1}{|A|}\mathrm{adj}A.$$

21．设

$$A=\frac{1}{2}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix},\quad I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

计算 $|A-\lambda I|$，其中 λ 是纯量．

22．证明，含有三个不共线点 $(a_i,b_i,c_i)(i=1,2,3)$ 的平面的方程是

$$\begin{vmatrix} x & y & z & 1 \\ a_1 & b_1 & c_1 & 1 \\ a_2 & b_2 & c_2 & 1 \\ a_3 & b_3 & c_3 & 1 \end{vmatrix} = 0.$$

23. 求通过点(8,－2,2)，(2,1,－4)与(2,4,－6)的平面的方程.

12.5 行列式与线性方程组

行列式的最重要的应用基于下面的定理，这个定理总结了正方矩阵、线性方程组以及行列式之间的相互关系.

定理 XI 设 A 是 $n \times n$ 矩阵，则下列各款是等价的：

(a) A 是奇异的(即 A 没有逆矩阵).

(b) $|A|=0$.

(c) n 个未知数 n 个方程的齐次方程组 $A\mathbf{x}=0$ 有非平凡解(即异于 $x_1=x_2=\cdots=x_n=0$ 的解).

(d) A 的行线性相关.

(e) A 的列线性相关.

证明 前四款的等价性由 12.2 节的定理 VI 以及 11.3 节的定理 VI 与定理 VII 得出. (c) 款与其他各款的等价性，可从如下的等价条款得以证实：

A 有线性相关的列；

A 的转置 A' 有线性相关的行；

$|A'|=0$；

$|A|=0$ (12.3 节的定理 X).

练　　习

24. 试求，当 λ 为何值时，方程组

$$\begin{aligned} -2x_1 + x_2 + x_3 &= \lambda x_1 \\ x_1 - 2x_2 + x_3 &= \lambda x_2 \\ x_1 + x_2 - 2x_3 &= \lambda x_3 \end{aligned}$$

有非平凡解？对每一个这样的 λ，求方程组的通解。（这一计算题是研究振动问题的开端。）

25. 求判定 x,y,z 的齐次方程组

$$x+a^2y+a(b+c)z=0$$
$$x+b^2y+b(c+a)z=0$$
$$x+c^2y+c(a+b)z=0$$

有非平凡解的行列式条件。把使非平凡解存在的 a,b,c 可能的值分类。就 $a=b,a\neq c$ 的情况，求出方程组的所有解。

26. 假定 y_1 与 y_2 是变量 t 的可微函数。证明，若 y_1 与 y_2 线性相关，则对 t 的任意值，

$$W(y_1,y_2)=\begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix}$$

都是零。（此处 y' 表示 y 的导数。）函数 $W(y_1,y_2)$ 称为 y_1 与 y_2 的朗斯基行列式，它在微分方程理论中十分重要。

27. 利用练习 26 的结果证明，函数

$$y_1(t)=\sin\ \omega t,\qquad y_2(t)=\cos\ \omega t\qquad (\omega\neq 0)$$

线性无关。

28. 函数 y_1,y_2 与 y_3 的朗斯基行列式定义为

$$W(y_1,y_2,y_3)=\begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix}.$$

利用这个行列式证明，函数

$$y_1(t)=e^t,\qquad y_2(t)=e^{-t},\qquad y_3(t)=1$$

线性无关。

第十三章　特征值问题

13.1　引论

设 A 是 $n\times n$ 矩阵。向量方程

$$A\mathbf{x}=\lambda\mathbf{x} \tag{1}$$

对所有纯量 λ 都有平凡解 $\mathbf{x}=\mathbf{0}$. 而对某些 λ 值，(1) 有非平凡解 $\mathbf{x}\neq\mathbf{0}$，这也是可能的. 任意一个这样的 λ 称为 A 的特征值*，而方程(1)对应的非零解 $\mathbf{x}$ 称 A 的对应于特征值 λ 的特征向量. A 的所有特征值的集合称为 A 的谱.

这一章就来研究矩阵论中的特征值问题及其应用. 这里先举几个例子作为开头.

例 1　若 A 是旋转的矩阵**，则 A 有特征值 1. 此外，A 的谱不能含有除 $\lambda=1$ 或 $\lambda=-1$ 以外的任何实数.

证明　旋转轴上任意一点都被该旋转固定，因此，方程 $A\mathbf{x}=1\,\mathbf{x}$ 有非平凡解. 因为旋转是等矩变换，对所有 $\mathbf{x}$，$A\mathbf{x}$ 的长度与 $\mathbf{x}$ 的长度相等. 特别，如果 $\mathbf{x}$ 是 A 的对应于特征值 λ 的特征向量，则

$$\|\mathbf{x}\|=\|A\mathbf{x}\|=\|\lambda\mathbf{x}\|=|\lambda|\,\|\mathbf{x}\|.$$

因为 $\|\mathbf{x}\|\neq 0$，推出 $|\lambda|=1$，这就是所要证明的. 特征值 -1 出现于 A 表示转 180° 的旋转时.

---

* 也可用“特征根”这一术语.

** 似应改为：“若 A 是三维空间的旋转的矩阵，……”因为二维空间的旋转只有旋转中心而没有旋转轴，若 $A\neq I$，则除原点外，没有任何固定点，故 A 没有特征值 1——译者注.

例2 矩阵 A 是奇异矩阵当且仅当 $\lambda=0$ 是它的特征值之一.

证明 据12.5节的定理XI，A 是奇异矩阵当且仅当方程组 $A\mathbf{x}=\mathbf{0}$ 有非平凡解，亦即，当且仅当 $A\mathbf{x}=0\,\mathbf{x}$ 有非平凡解. 证毕.

接着，我们再举两个例子以说明在应用数学中特征值问题是如何出现的. 其中第一个例子与分子振动理论中所碰到的问题有某些相似之处.

例3 振动问题

把一条长度为 $3a$ 的弹簧放在光滑的水平桌面上，让其处于自然状态，然后固定它的两端. 再把两个质量都是 m 的质点加在这条弹簧的三等分点上. 当这一系统在弹簧所在的直线上振动时，讨论这两个质点的运动（弹簧的质量忽略不计).

根据虎克定律，如果自然长度为 a 的弹簧被拉长或压缩了 s，弹簧产生的恢复力与 s 成正比. 比例常数 k 称为该弹簧的弹性系数，它由下式给出：

$$k=\frac{\text{弹簧的弹性模量}}{\text{弹簧的自然长度}}.$$

在我们的例子中，三段弹簧具有相同的弹性系数 k.

该系统的运动由微分方程*

$$\begin{aligned} m\ddot{x}_1&=F_1-F_2=k(x_2-x_1)-kx_1,\\ m\ddot{x}_2&=F_3-F_2=-kx_2-k(x_2-x_1) \end{aligned} \tag{2}$$

所支配. 这里按惯例以 $\ddot{x}$ 代替 $\frac{d^2x}{dt^2}$.

注意到在图79中，若 x_2 是正的，则右边的弹簧处于压缩状态，于是 F_3 是负的：$F_3=-kx_2$.

---

* 这里，我们按"力＝质量×加速度"列出运动方程. 对更复杂的动力学系统，或者为了理论的探讨，人们更喜欢采用拉格朗日运动方程（参看16.4节).

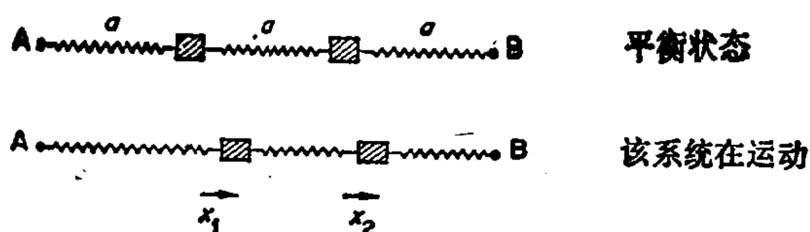


图 79 振动的弹簧

从方程(2) 我们得到

$$\ddot{x}_1=q(-2x_1+x_2)$$
$$\ddot{x}_2=q(x_1-2x_2), \tag{3}$$

其中 $q=\dfrac{k}{m}$.注意到方程(3)可以写成一个向量方程

$$\ddot{\mathbf{x}}=A\mathbf{x},$$

其中

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix},\quad A=\begin{pmatrix}-2q & q\\ q & -2q\end{pmatrix},$$

而向量的导数是按分量分别求导.

我们现在问：这一动力学系统是否可能有一种运动方式，使两个质点以同一频率振动？让我们来求方程(3) 的形如

$$x_1=c_1\sin(\omega t+\alpha),\quad x_2=c_2\sin(\omega t+\alpha) \tag{4}$$

的非平凡解，这里 c_1，c_2，ω 及 α 都是与 t 无关的常数.

把(4) 式代入(3)，我们得到

$$A\mathbf{c}=-\omega^2\mathbf{c},\qquad \mathbf{c}=\begin{pmatrix}c_1\\c_2\end{pmatrix}. \tag{5}$$

方程(5) 的每一非平凡解 $\mathbf{c}$ 都是 A 的特征向量，其所对应的特征值为 $\lambda=-\omega^2$.

至此，我们把问题归结为一个特征值问题：A 的每一个负特征值 λ，对应这一系统一个可能的振动方式，使两个质点以同一角频

率 $\omega=\sqrt{-\lambda}$ 作振动．这样的振动方式称为该系统的正规振动方式．(我们没有必要关心如何解释 A 的非负特征值，因为 A 根本没有非负特征值！)

在 13.2 节的例 9 以及 13.3 节中，我们还要回过头来解这个个振动问题．

取自应用数学的第二个例子涉及的是边界值问题中的特征值，边界值问题是指有关带已知边界条件的微分方程的一些问题．

例 4 边界值问题

试求，当 k 为何值时，微分方程

$$\frac{d^2y}{dx^2}+ky=0 \tag{6}$$

有满足边界条件

$$y(0)=0 \quad 与 \quad y(l)=0 \tag{7}$$

的非平凡解？

为了解决这个问题，我们分三种情况讨论：$k=0$，$k<0$ 以及 $k>0$．

情况 1 $k=0$

方程(6) 的通解是

$$y=c_1x+c_2.$$

如果加上边界条件 (7)，我们就得到 $c_1=c_2=0$．因此，当 $k=0$ 时，方程(6)没有适合条件的非平凡解．

情况 2 $k<0$

置 $k=-r^2$．方程(6)的通解是

$$y=c_1e^{rx}+c_2e^{-rx}.$$

加上边界条件，又使 $c_1=c_2=0$．

情况 3 $k>0$

置 $k=r^2$．方程(6) 的通解是

$$y=c_1\sin rx+c_2\cos rx.$$

从 $y(0)=0$ 得到 $c_2=0$；条件 $y(l)=0$ 使得 $c_1\sin rl=0$. 所以，只要 $\sin rl=0$，亦即只要

$$r=\pm\frac{\pi}{l},\ \pm\frac{2\pi}{l},\ \pm\frac{3\pi}{l},\ \cdots,$$

函数

$$y=c\sin rx$$

就是方程(6) 满足所给边界条件的非平凡解. 解毕.

我们来解释一下，为什么上面的问题可以看成特征值问题. 微分方程(6) 可以写成

$$D^2y=\lambda y, \tag{8}$$

这里 D 代替了变换 $\frac{d}{dx}$，而 k 换成了 $-\lambda$. 于是，很自然地把满足边界条件的 λ 值称为该边界值问题的特征值，对应的非平凡解称为特征向量或特征函数.

这样，上述问题中特征值* 为 $\lambda=-k$，其中

$$k=\frac{\pi^2}{l^2},\ \frac{4\pi^2}{l^2},\ \frac{9\pi^2}{l^2},\ \cdots,\ \frac{n^2\pi^2}{l^2},\cdots,$$

对应的特征函数为

$$y_n(x)=c_n\sin\frac{n\pi}{l}x\quad(n=1,2,\cdots).$$

把一根长度为 l 的弦两端固定. 这根弦的振动是由一个偏微分方程所支配，这个方程可以化为带边界条件 (7)的常微分方程(6). 它的解 $y_n(x)$表示了这根弦的正规振动方式 (图80).

上例说明了在研究带边界条件的微分方程时，会很自然地产生出特征值问题. 由于物理学与应用数学的大量问题(如波的运动，热流，量子理论) 都基于这样的微分方程，特征值问题的重要

* 有些教科书把数 k (而不是 $-k$)称为该问题的特征值.

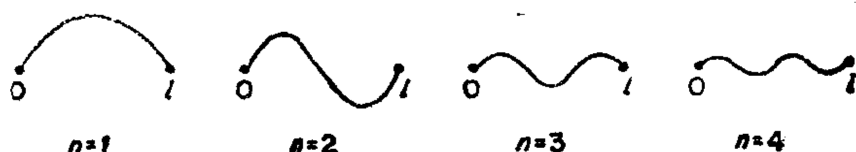


图 80 弦的正规振动方式

性就显而易见了．作了上述说明之后，我们将离开这些应用问题，把本章剩下的篇幅只用于讨论矩阵论中出现的特征值问题．

练　　习

1．证明，反射矩阵的特征值为 1 与 − 1．

2．矩阵

$$A=\frac{1}{9}\begin{pmatrix} 7 & 4 & -4 \\ 4 & 1 & 8 \\ -4 & 8 & 1 \end{pmatrix}$$

表示的是旋转或反射．请判断之，然后再求旋转轴或反射平面．

3．证明，正交射影的矩阵是奇异的．

4．假定矩阵 A 有一个特征值是 λ，对应的特征向量为 $\mathbf{x}$。证明，矩阵 qA 有一个特征值是 $q\lambda$，对应的特征向量为 $\mathbf{x}$。

5．证明，若矩阵 A 有一个特征值是 λ，则 A^r 有一个特征值是 λ^r， 而矩阵

$$c_0I+c_1A+\cdots+c_rA^r$$

有一个特征值是 $c_0+c_1\lambda+\cdots+c_r\lambda^r$．

6．证明，切变矩阵

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

仅有的特征值是 $\lambda=1$。

7．正方矩阵 A 满足 $A^2=A$(正交射影的矩阵就是一例)．证明，A不可能有异于 0 与 1 的特征值．

13.2　矩阵的特征方程

设 $A=(a_{ij})$ 是 $n\times n$ 矩阵，λ 是 A 的一个特征值，则方程

$Ax=\lambda x$ 有非平凡解．这等价于说方程

$$(A-\lambda I)x=0 \tag{9}$$

有非平凡解(这里 I 是 $n\times n$ 恒等矩阵)．根据12.5节的定理XI，方程(9)有非平凡解当且仅当 $A-\lambda I$ 的行列式值为零．所以，我们得到如下的重要定理．

定理 I λ 是 A 的特征值当且仅当

$$|A-\lambda I|=0. \tag{10}$$

注意到

$$f(\lambda)=|A-\lambda I|=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}-\lambda \end{vmatrix} \tag{11}$$

是 λ 的 n 次多项式，我们把 $f(\lambda)$ 称为 A 的特征多项式，把方程(10)称为 A 的特征方程．A 的特征值正好是 A 的特征方程的根．

把(11)式展开，我们得到

$$f(\lambda)=(-1)^n\lambda^n+(-1)^{n-1}(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+c. \tag{12}$$

系数 $(a_{11}+a_{22}+\cdots+a_{nn})$ 称为矩阵 A 的迹．常数项是

$$c=f(0)=|A-0I|=|A|.$$

A 的特征多项式中其余系数没有简单的表达式．

每一个实系数或复系数的 n 次多项式至少有一个根，但其不同根的总数不能超过 n 个(其中某些根或者全部根可能是复数)．所以，一个实或复的 $n\times n$ 矩阵至少有一个特征值，但其特征值的总数不能超过 n 个．其中某些或全部特征值可能是复数*(参看例

* 严格地说，实数也是复数，所以，一个实或复的正方矩阵的特征值全是复数．但本段中“复数”一词应理解为虚部不等于零的复数——译者注。

8). 作为直接推论,我们有如下一些结果.

定理 II 设A是实或复的正方矩阵,则存在一个(实或复的)纯量λ,使方程组$A\mathbf{x}=\lambda\mathbf{x}$有非平凡解.

可以证明,当A是实矩阵,λ是实特征值时,$\mathbf{x}$有非平凡的实解(参看练习 15).

我们注意到,对一个固定的λ,若$A\mathbf{x}=\lambda\mathbf{x}$,$A\mathbf{y}=\lambda\mathbf{y}$,则

$$A(c\mathbf{x}+d\mathbf{y})=cA\mathbf{x}+dA\mathbf{y}=c\lambda\mathbf{x}+d\lambda\mathbf{y}=\lambda(c\mathbf{x}+d\mathbf{y}).$$

因此,$\mathbf{x}$与$\mathbf{y}$的每一线性组合,只要不是零向量,就也是A的特征向量,对应的特征值还是λ. 这就证明了如下的定理.

定理 III 设λ是$n\times n$矩阵A的一个特征值, 则A的对应于λ的所有特征向量的集合再添上零向量,便组成n维列向量空间的一个子空间$\mathscr{S}_\lambda$.

例 5 求矩阵

$$A=\begin{pmatrix}-2 & 1\\ 1 & -2\end{pmatrix}$$

的特征值及其对应的特征向量.

特征方程

$$\begin{vmatrix}-2-\lambda & 1\\ 1 & -2-\lambda\end{vmatrix}=\lambda^2+4\lambda+3=0$$

的根为$\lambda=-1$与$\lambda=-3$.

我们来求对应的特征向量. 对于$\lambda=-1$,方程$A\mathbf{x}=\lambda\mathbf{x}$成为

$$\begin{aligned}-2x_1+\ x_2&=-x_1,\\ x_1-2x_2&=-x_2,\end{aligned}$$

它的通解是$x_1=x_2$. 因此,向量$c\begin{pmatrix}1\\1\end{pmatrix}$(其中$c\neq0$)都是$A$的对应于特征值$-1$的特征向量. 对于$\lambda=-3$,方程$A\mathbf{x}=\lambda\mathbf{x}$成为

$$-2x_1+x_2=-3x_1,$$

$$x_1-2x_2=-3x_2,$$

它的通解是 $x_1=-x_2$. 因此，向量 $c\begin{pmatrix}-1\\1\end{pmatrix}$（其中 $c\neq0$）都是 A 的对应于特征值 -3 的特征向量. 我们可以把上述结果总结如下：子空间 $\mathscr{S}_{-1}$ 与 $\mathscr{S}_{-3}$ 分别由向量 $\begin{pmatrix}1\\1\end{pmatrix}$ 与 $\begin{pmatrix}-1\\1\end{pmatrix}$ 生成.

例 6 求矩阵

$$A=\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}$$

的特征值及其对应的特征向量.

A 的特征方程是

$$|A-\lambda I|=-(\lambda-3)^2(\lambda+6)=0.$$

所以，A 有两个特征值：$\lambda=3$ 与 $\lambda=-6$. 对于 $\lambda=3$，方程 $A\mathbf{x}=\lambda\mathbf{x}$ 成为

$$\begin{aligned}-x_1+2x_2+4x_3&=3x_1,\\ 2x_1+2x_2-2x_3&=3x_2,\\ 4x_1-2x_2-x_3&=3x_3.\end{aligned}$$

其中每个方程都可化为

$$2x_1-x_2-2x_3=0,$$

所以，该方程组的通解是

$$x_1=c,\quad x_3=d,\quad x_2=2(c-d)\ (c \text{ 与 } d \text{ 任意}).$$

A 的对应于 $\lambda=3$ 的特征向量都可表为

$$\begin{pmatrix}c\\ 2c-2d\\ d\end{pmatrix}=c\begin{pmatrix}1\\2\\0\end{pmatrix}+d\begin{pmatrix}0\\-2\\1\end{pmatrix}=c\mathbf{v}_1+d\mathbf{v}_2$$

的形式，其中 $\mathbf{v}_1=\begin{pmatrix}1\\2\\0\end{pmatrix}$，$\mathbf{v}_2=\begin{pmatrix}0\\-2\\1\end{pmatrix}$. 因此，$\mathscr{S}_3$ 是二维空间，由 $\mathbf{v}_1$ 与 $\mathbf{v}_2$ 生成. 同理可求得 $\mathscr{S}_{-6}$ 由向量

$$\mathbf{v}_3=\begin{pmatrix}-2\\1\\2\end{pmatrix}$$

生成.

例 7　求矩阵

$$A=\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$$

的特征值及其对应的特征向量.

因为

$$|A-\lambda I|=(\lambda-1)^2,$$

所以 A 仅有一个特征值 $\lambda=1$. 对应的特征向量是方程组

$$\begin{aligned}x_1+x_2&=x_1\\x_2&=x_2\end{aligned}$$

的非平凡解. 因此 $x_2=0$, x_1 任意. $\mathscr{S}_1$ 由 $\begin{pmatrix}1\\0\end{pmatrix}$ 生成. 这个例子从几何上可以这样解释：矩阵 A 表示在 x_1 方向的切变. 这个切变把 x_1 轴上的向量都映到自身，所以它有特征值 1；而其他任何向量都被此切变变换改变了方向.

例 8

矩阵

$$A=\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$$

的特征方程为 $|A-\lambda I|=\lambda^2+1=0$. 矩阵 A 没有实特征值. 这

是毫不奇怪的，因为我们可以认出，在一个直角坐标系中，A表示转 90° 的旋转，所有非零向量经过这个旋转都改变了方向.

但是，如果我们允许复数出现，则A有两个特征值：$\lambda=i$ 与 $\lambda=-i$. 对应的特征向量子空间 $\mathscr{S}_i$ 与 $\mathscr{S}_{-i}$ 分别由$\begin{pmatrix}1\\-i\end{pmatrix}$与$\begin{pmatrix}1\\i\end{pmatrix}$生成. A没有各分量全是实数的特征向量.

这个例子说明了交代清楚矩阵A的元素及其特征向量的分量取自什么数系是很重要的事. 一个实矩阵可能只有复的特征向量. 但是，如果一个实矩阵有一个实特征值，那么，该特征值对应的特征向量可以取为各分量全是实数的向量.

我们注意到，在以上各例中，矩阵A的对应于不同特征值 λ^1 与 λ_2 的特征向量 $\mathbf{v}_1$ 与 $\mathbf{v}_2$ 是线性无关的. 这一结论总是成立的. 因为若假定有线性关系

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2=\mathbf{0}, \tag{13}$$

则

$$A(c_1\mathbf{v}_1+c_2\mathbf{v}_2)=\mathbf{0},$$

所以

$$c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2=\mathbf{0}. \tag{14}$$

从(14)式减去(13)式的 λ_2 倍，便得到

$$(\lambda_1-\lambda_2)c_1\mathbf{v}_1=\mathbf{0},$$

由于 $\lambda_1\neq\lambda_2$ 且 $\mathbf{v}_1\neq\mathbf{0}$，得出 $c_1=0$. 再从关系式(13)得到 $c_2=0$. 所以，$\mathbf{v}_1$ 与 $\mathbf{v}_2$ 之间仅有的关系是平凡关系.

上面的论证加以推广便导出如下的定理.

定理 IV 矩阵 A 的对应于不同特征值 $\lambda_1,\cdots,\lambda_r$ 的特征向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 是线性无关的.

证明 用归纳法. $r=2$ 的情况上面已证过. 假设定理对 $r<k$ 都已证明. 现在设 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 是 A 的对应不同特征值 $\lambda_1,\cdots,\lambda_k$

的特征向量．考虑线性关系

$$c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k=\mathbf{0}. \tag{15}$$

此时，

$$A(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=\mathbf{0},$$

所以

$$c_1\lambda_1\mathbf{v}_1+\cdots+c_k\lambda_k\mathbf{v}_k=\mathbf{0}. \tag{16}$$

从(16)式减去(15)式的 λ_k 倍，便得到

$$(\lambda_1-\lambda_k)c_1\mathbf{v}_1+(\lambda_2-\lambda_k)c_2\mathbf{v}_2+\cdots+(\lambda_{k-1}-\lambda_k)c_{k-1}\mathbf{v}_{k-1}=\mathbf{0}. \tag{17}$$

由归纳假设，$\mathbf{v}_1,\cdots,\mathbf{v}_{k-1}$ 线性无关，因而

$$(\lambda_i-\lambda_k)c_i=0 \quad (i=1,\cdots,k-1).$$

由于特征值 $\lambda_1,\cdots,\lambda_k$ 互不相同，推及

$$c_i=0 \quad (i=1,\cdots,k-1).$$

再由关系式(15)得到 $c_k=0$．于是，定理对 $r=k$ 也成立．证毕．

例 9　振动问题中的正规方式

我们回到 13.1 节的例 3．矩阵

$$A=\begin{pmatrix}-2q & q\\ q & -2q\end{pmatrix}$$

有两个特征值：$\lambda=q$ 与 $\lambda=-3q$．$\mathscr{S}_{-q}$ 与 $\mathscr{S}_{-3q}$ 分别由向量 $\begin{pmatrix}1\\1\end{pmatrix}$ 与 $\begin{pmatrix}-1\\1\end{pmatrix}$ 生成（同例 5 比较，在例 5 中我们考虑了 $q=1$ 的情况）．所以有两类正规振动方式，它们由 13.1 节的(4)、(5)式定义，其中

(i)　$\omega^2=q$，$\begin{pmatrix}c_1\\c_2\end{pmatrix}=c\begin{pmatrix}1\\1\end{pmatrix}$（$c$ 任意）；

(ii) $\omega^2=3q, \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}=d\begin{pmatrix} -1 \\ 1 \end{pmatrix}$($d$ 任意).

这些振动分别由下列两组方程所描述：

(i) $x_1=c\sin(\sqrt{q}t+\alpha)$, $x_2=c\sin(\sqrt{q}t+\alpha)$;

(ii) $x_1=-d\sin(\sqrt{3q}t+\beta)$, $x_2=d\sin(\sqrt{3q}t+\beta)$.

正规振动方式的振幅可以任取．当此系统以正规振动方式振动时，两个质点同时达到其偏离平衡位置的极大位移(图 81)．两类正规振动方式的频率之比为 $1:\sqrt{3}$.



图 81 弹簧振动的正规方式

要观察正规振动方式，可以把两个质点移离其平衡位置相同的距离，或者移向同侧(方式(i))，或者移向异侧(方式(ii))，然后同时放开这两个质点，使这一系统运动．如果以其他方式使这个系统运动，进行的就不是简谐振动．但从这一振动问题通解的形式可以明显看出，该运动仍是上述正规方式的线性组合，下节我们就要进行这一讨论.

练 习

8．设 $A=(a_{ij})$ 与 $B=(b_{ij})$ 都是 3×3 矩阵．求 AB 的迹，并证明 AB 的迹与 BA 的迹相同．试推广之．举数字的例子加以说明．

9．把例 6 与例 8 的解答补充完整．

10．设 A 是反射的矩阵，求它的特征向量空间 $\mathscr{S}_1$ 与 $\mathscr{S}_{-1}$ 的维数．

11．求下列矩阵的特征值及其对应的特征向量空间：

(i) $\begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}$; (ii) $\begin{pmatrix} 3 & 4 \\ -1 & -1 \end{pmatrix}$; (iii) $\begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ -2 & 0 & 4 \end{pmatrix}$;

(iv) $\begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix}$； (v) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

12．一个力学系统的运动服从于方程

$$\ddot{x}_1 = -5x_1 + 2x_2,$$

$$\ddot{x}_2 = 2x_1 - 2x_2.$$

描述其正规振动方式，并证明两类正规方式的频率比为 $1:\sqrt{6}$.

13．证明，如果 λ 是矩阵 A 的特征值，$\mathbf{x}$ 为所对应的特征向量，则 $\bar{\lambda}$ 是矩阵 $\bar{A}$ 的特征值，且 $\bar{\mathbf{x}}$ 为所对应的特征向量(矩阵 $\bar{A}$ 表示 A 的共轭矩阵).

14．证明，矩阵

$$A = \begin{pmatrix} 2+i & -i \\ -i & 2+i \end{pmatrix}$$

的特征值为 $2+2i$ 与 2. 求对应的特征向量空间. 写出 A 的共轭矩阵 $\bar{A}$，并求它的特征值与对应的特征向量.

15．证明，如果实矩阵 A 有实特征值 λ，则 λ 对应的特征向量可取为实向量.

[提示：把 12.5 节的定理 XI 用于矩阵 $A-\lambda I$.]

13.3 坐标变换；矩阵的相似性

振动问题的解决

由于坐标系选择适当，常常会使一个问题变得易于解决. 我们以 13.1 节例 3 与 13.2 节例 9 所涉及的振动问题来说明这一点. 该问题是要解一个微分方程组，这方程组的向量形式是

$$\ddot{\mathbf{x}} = A\mathbf{x}, \quad \text{其中 } A = \begin{pmatrix} -2q & q \\ q & -2q \end{pmatrix}. \tag{18}$$

我们引进新的坐标 y_1, y_2，它们与旧坐标有线性关系

$$\mathbf{x} = C\mathbf{y},$$

完整地写出来，是

$$\begin{aligned} x_1 &= c_{11}y_1 + c_{12}y_2, \\ x_2 &= c_{21}y_1 + c_{22}y_2. \end{aligned} \tag{19}$$

C 取为非奇异的，使新坐标也能用旧坐标表出：

$$\mathbf{y} = C^{-1}\mathbf{x}. \tag{20}$$

把方程组(19)代入(18)，我们得到

$$C\ddot{\mathbf{y}} = AC\mathbf{y},$$

所以

$$\ddot{\mathbf{y}} = (C^{-1}AC)\mathbf{y}. \tag{21}$$

关于新坐标的方程(21)等价于关于旧坐标的方程(18)．我们现在要说明的是，如果 C 选择适当，方程组(21)的解可以一眼看出：

置

$$C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

则

$$C^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

$$AC = \begin{pmatrix} -2q & q \\ q & -2q \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -q & 3q \\ -q & -3q \end{pmatrix}, \tag{22}$$

$$C^{-1}AC = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -q & 3q \\ -q & -3q \end{pmatrix} = \begin{pmatrix} -q & 0 \\ 0 & -3q \end{pmatrix}. \tag{23}$$

代入方程组(21)，便得到

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} = \begin{pmatrix} -q & 0 \\ 0 & -3q \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{24}$$

亦即

$$\begin{aligned} \ddot{y}_1 + qy_1 &= 0, \\ \ddot{y}_2 + 3qy_2 &= 0. \end{aligned} \tag{25}$$

方程组(25)的通解是

$$y_1=c\sin(\sqrt{q}\,t+\alpha),\quad y_2=d\sin(\sqrt{3q}\,t+\beta).$$

由(19),

$$x_1=y_1-y_2,\quad x_2=y_1+y_2.$$

所以,我们得到这个问题的通解(用旧坐标写出)是

$$x_1=c\sin(\sqrt{q}\,t+\alpha)-d\sin(\sqrt{3q}\,t+\beta),$$
$$x_2=c\sin(\sqrt{q}\,t+\alpha)+d\sin(\sqrt{3q}\,t+\beta).$$

它们是正规振动方式的线性组合. 常数 c,d,α,β 取决于两个质点的初始位置和速度.

坐标 y_1,y_2 称为这一系统的正规坐标,因为这组坐标指示出了正规振动方式;分别令 $y_2=0$, $y_1\neq0$ 及 $y_1=0$, $y_2\neq0$, 就得到两类正规方式.

上述振动问题的解决依赖于选择矩阵 C, 使 $C^{-1}AC$ 是对角矩阵. 它使我们得到一个微分方程组,其中每个方程只含一个变量,于是,可以用观察法解这个方程组.

显然,上述方法可相当广泛地应用于解 $x_1,\cdots,x_n$ 的方程组 $\dot{\mathbf{x}}=A\mathbf{x}$,这里 A 是 $n\times n$ 矩阵,只要我们能够找到一个非奇异的矩阵 C,使 $C^{-1}AC$ 是对角矩阵. 现在我们要讨论什么时候和怎样才能找到这样的矩阵 C.

定义 I 设 A 与 B 是 $n\times n$ 矩阵. 如果存在一个非奇异矩阵 C 使 $B=C^{-1}AC$,我们就说 B 相似于 A.

相似关系是对称关系:如果 B 相似于 A,则 A 相似于 B.

证明 $B=C^{-1}AC$ 意味着 $A=C_1^{-1}BC_1$,其中 $C_1=C^{-1}$.

因此,我们可以说 A 与 B 是相似矩阵,绝对不会有歧义. 相似关系还具有自反性与传递性(参看练习 21). 所以,相似关系是一个等价关系(参看附录).

用相似关系的语言,可以把我们的问题复述如下:矩阵 A 相似于对角矩阵吗? 如果它相似于对角矩阵,怎样找到矩阵 C,使

$C^{-1}AC$ 是对角矩阵？答案如下.

定理 V 设 A 与 C 是 $n\times n$ 矩阵，C 非奇异. 则，$C^{-1}AC$ 是对角阵，亦即

$$C^{-1}AC=\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix},$$

当且仅当 C 的第一列，第二列，$\cdots$，第 n 列都是 A 的特征向量，且对应的特征值分别为 $\lambda_1,\cdots,\lambda_n$.

证明 先假定 C 的列

$$\mathbf{c}_1=\begin{pmatrix}c_{11}\\ c_{21}\\ \vdots\\ c_{n1}\end{pmatrix},\quad \mathbf{c}_2=\begin{pmatrix}c_{12}\\ c_{22}\\ \vdots\\ c_{n2}\end{pmatrix},\quad \cdots,\mathbf{c}_n=\begin{pmatrix}c_{1n}\\ c_{2n}\\ \vdots\\ c_{nn}\end{pmatrix}$$

都是 A 的特征向量，且对应的特征值分别为 $\lambda_1,\cdots,\lambda_n$. 则

$$\begin{aligned}AC&=A(\mathbf{c}_1\ \ \mathbf{c}_2\cdots\mathbf{c}_n)=(A\mathbf{c}_1\ \ A\mathbf{c}_2\cdots A\mathbf{c}_n)\text{（练习 16）}\\ &=(\lambda_1\mathbf{c}_1\ \ \lambda_2\mathbf{c}_2\cdots\lambda_n\mathbf{c}_n)\\ &=\begin{pmatrix}\lambda_1c_{11} & \lambda_2c_{12} & \cdots & \lambda_nc_{1n}\\ \lambda_1c_{21} & \lambda_2c_{22} & \cdots & \lambda_nc_{2n}\\ \vdots & \vdots & & \vdots\\ \lambda_1c_{n1} & \lambda_2c_{n2} & \cdots & \lambda_nc_{nn}\end{pmatrix}\\ &=\begin{pmatrix}c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & & \vdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix}\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}\\ &=CD,\end{aligned}$$

其中 D 为 ii-元素是 λ_i 的对角矩阵．由此可见，

$$C^{-1}AC=D.$$

反过来，若已知 $C^{-1}AC=D$（对角矩阵），则 $AC=CD$．根据上面的计算，$Ac_i=\lambda_i c_i$ $(i=1,\cdots,n)$．证毕．

定理V的推论　$n\times n$ 矩阵 A 相似于对角矩阵当且仅当 A 有 n 个线性无关的特征向量．

证明　由定理V，A 相似于对角矩阵当且仅当存在一个非奇异矩阵，它的 n 个列都是 A 的特征向量．也就是说，存在一个矩阵，它的 n 个列组成 A 的线性无关特征向量的集合．证毕．

如果矩阵 A 相似于对角矩阵，我们就说 A 可对角化．所谓对角化过程，就是找一个矩阵 C，使 $C^{-1}AC$ 是对角矩阵．

现在我们已经很清楚地知道，在本节开头解振动问题时，我们是如何选中矩阵 C 的：我们把 C 的列取为 A 的线性无关的特征向量．

例 10　检验一下 13.2 节各例题所讨论的矩阵，看其中哪些可对角化（即相似于一个对角矩阵）．

在例 5 中，矩阵

$$A=\begin{pmatrix}-2 & 1\\ 1 & -2\end{pmatrix}$$

有两个线性无关的向量．使 $C^{-1}AC$ 是对角矩阵的 C 可以有各种不同的选择法．相当一般地说，如果

$$C=\begin{pmatrix}c & -d\\ c & d\end{pmatrix}\quad(c\neq0,d\neq0),$$

则

$$C^{-1}AC=\begin{pmatrix}-1 & 0\\ 0 & -3\end{pmatrix},$$

如果

$$C=\begin{pmatrix}-d & c\\ d & c\end{pmatrix}\quad (c\neq 0, d\neq 0),$$

则

$$C^{-1}AC=\begin{pmatrix}-3 & 0\\ 0 & -1\end{pmatrix}.$$

根据定理Ⅴ，上述结果是显然的.

在例 6 中：矩阵

$$C=\begin{pmatrix}1 & 0 & -2\\ 2 & -2 & 1\\ 0 & 1 & 2\end{pmatrix}$$

是非奇异的，且

$$C^{-1}\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}C=\begin{pmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & -6\end{pmatrix}.$$

当然，C还可以有其他取法.

在例 7 中：因为$\mathscr{S}_1$是一维的，所以不可能找到A的两个线性无关的特征向量. 切变矩阵不是可对角化的.

在例 8 中：该旋转矩阵不能被一个实矩阵对角化. 但是，

$$\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}^{-1}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}=\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}.$$

定理 Ⅵ 设 $n\times n$ 矩阵 A 有 n 个不同的特征值，则 A 相似于对角矩阵.

证明 据 13.2 节的定理 Ⅳ，A的对应于不同特征值 $\lambda_1,\cdots,\lambda_n$ 的特征向量 $\mathbf{x}_1,\cdots,\mathbf{x}_n$ 是线性无关的. 设C是 $\mathbf{x}_1,\cdots,\mathbf{x}_n$ 为列的 $n\times n$ 矩阵，则 A是非奇异的. 根据定理 Ⅴ，$C^{-1}AC$ 是对角矩阵. 证毕.

一个 $n\times n$ 矩阵的不同特征值若少于 n 个，则它可能相似于对角矩阵，也可能不相似于对角矩阵. 这两种可能性在例 10 中都

已举例说明，即例 6 与例 7 出现的矩阵．其中前一个例子是对称矩阵．在 15.3 节中我们将证明，所有对称矩阵都可对角化．

从下面的定理可以推出一些重要的结果．

定理 VII 相似的矩阵有相同的特征多项式．

证明

$$
\begin{aligned}
|C^{-1}AC-\lambda I| &= |C^{-1}AC-\lambda C^{-1}C| \\
&= |C^{-1}(A-\lambda I)C| \\
&= |C^{-1}||A-\lambda I||C| = |A-\lambda I|,
\end{aligned}
$$

其中最后一个等号是由于 $|C^{-1}||C|=|C^{-1}C|=|I|=1$．由此可见，$C^{-1}AC$ 与 A 有相同的特征多项式．证毕．

定理 VII 的推论 相似的矩阵有相同的迹与行列式值．

证明 由 13.2 节的 (12) 式，迹与行列式值都作为特征多项式的系数出现．

在 14.2 节的例 4 中，我们将给出该结果的一个应用．

练 习

16．证明定理 V 的证明中所用到的公式

$$A(\mathbf{c}_1\ \mathbf{c}_2\cdots\mathbf{c}_n)=(A\mathbf{c}_1\ A\mathbf{c}_2\cdots A\mathbf{c}_n),$$

并以如下实例作说明：

$$A=\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -1\\ 4 & -2 & -1\end{pmatrix},\quad C=\begin{pmatrix}1 & 0 & -2\\ 2 & -2 & 1\\ 0 & 1 & 2\end{pmatrix}.$$

17．判定，13.2 节练习 11 的矩阵中，哪些相似于对角矩阵？对每个可对角化矩阵 A，找到使 $C^{-1}AC$ 是对角矩阵的矩阵 C，并求出 $C^{-1}AC$。

18．设 A 是 13.2 节例 6 的矩阵．求矩阵 C，使

$$C^{-1}AC=\begin{pmatrix}3 & 0 & 0\\ 0 & -6 & 0\\ 0 & 0 & 3\end{pmatrix}.$$

19．证明，如果正文所讨论的弹簧不是按正规方式之一运动，则它的运动不是简谐振动．

[提示：证明，在作非正规运动时，两个质点的任何位置都不可能重复第二次。]

20. 求 13.2 节练习 12 的振动问题的通解。

21. 证明，相似关系是

(i) 自反的关系：矩阵 A 相似于自己。

(ii) 传递的关系：如果 A_1 相似于 A_2，A_2 相似于 A_3，则 A_1 相似于 A_3。

22. 证明，$(C^{-1}AC)^k = C^{-1}A^kC$。

23. 求对角矩阵

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

的 k 次幂。试推广之。

第十四章 正交矩阵

14.1 直角坐标系的变换：正交矩阵

在许多与几何有关的问题中，参照系都取为直角坐标系．我们常常需要把坐标系变换到一个新的直角坐标系，使手头的问题变得易于处理*．我们先看一个简单的二维例子．

例 1 在图 82 中画有一个以 O 为中心的椭圆，它的长轴与水平轴倾成 45°，长、短二轴的长度分别为 $2a$ 与 $2b$．求此椭圆在 x_1, x_2 坐标系中的方程．

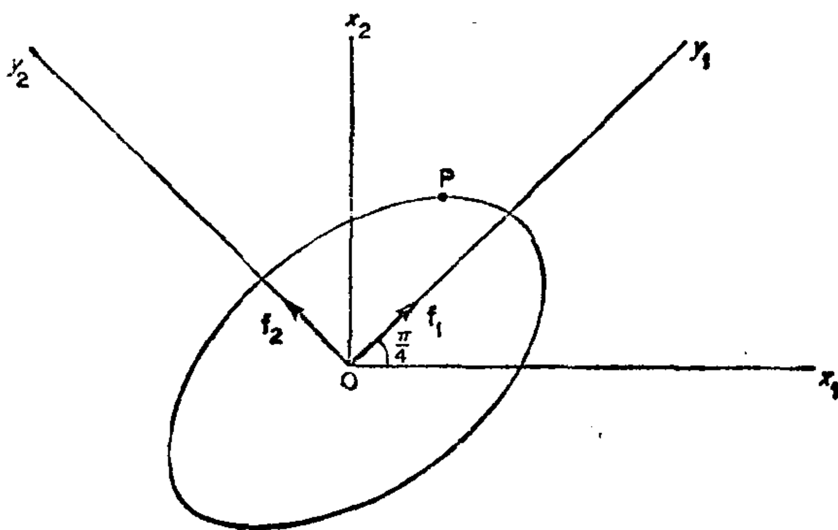


图 82 椭圆

沿着该椭圆长轴与短轴的方向引进"新"坐标系．在新坐标系中，椭圆的方程是

$$\frac{y_1^2}{a^2}+\frac{y_2^2}{b^2}=1 \tag{1}$$

* 在 13.3 节中我们已经看到选择适当的坐标系对解决振动问题所起的作用．但那个问题同直角坐标系毫不相干．

也就是说，如果一点 P 在新坐标系中的坐标为 (y_1, y_2)，则点 P 在椭圆上当且仅当它满足方程(1).

为了变回到"旧"坐标系，我们问：如果点 P 在旧坐标系中的坐标为 (x_1, x_2)，那么，x_1, x_2 与 y_1, y_2 有什么关系？这个问题用向量方法容易解决.

把新轴方向上的单位向量记为 $\mathbf{f}_1$ 与 $\mathbf{f}_2$. 在旧坐标系中把各向量用分量表出，我们有

$$\mathbf{f}_1=\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix},\quad \mathbf{f}_2=\begin{pmatrix}-\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix},\qquad \overrightarrow{OP}=\begin{pmatrix}x_1\\ x_2\end{pmatrix}.$$

但

$$\overrightarrow{OP}=y_1\mathbf{f}_1+y_2\mathbf{f}_2=y_1\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix}+y_2\begin{pmatrix}-\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix}, \tag{2}$$

所以

$$\begin{aligned} x_1&=\frac{1}{\sqrt{2}}y_1-\frac{1}{\sqrt{2}}y_2,\\ x_2&=\frac{1}{\sqrt{2}}y_1+\frac{1}{\sqrt{2}}y_2. \end{aligned} \tag{3}$$

从关系式(3)容易得到

$$y_1=\frac{1}{\sqrt{2}}(x_1+x_2),$$

$$y_2=\frac{1}{\sqrt{2}}(x_2-x_1).$$

代入方程(1)，我们求出了这个椭圆在旧坐标系中的方程：

$$\frac{\frac{1}{2}(x_1+x_2)^2}{a^2}+\frac{\frac{1}{2}(x_2-x_1)^2}{b^2}=1,$$

它可以写成

$$(a^2+b^2)(x_1^2+x_2^2)+2(b^2-a^2)x_1x_2=2a^2b^2. \tag{4}$$

这就解决了我们的问题．请注意，方程(4)的左边是 x_1, x_2 的二次型．

让我们从另一角度来看一看描述坐标变换的线性关系(3)．用矩阵记号可以把这一关系写成

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

或简记为

$$\mathbf{x}=C\mathbf{y}，其中 C=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{5}$$

比较(2)、(5)两式，我们看到，在旧坐标系中表示出来的向量 $\mathbf{f}_1$ 与 $\mathbf{f}_2$ 正好是矩阵 C 的列．

这种现象是相当普遍地存在的，所以可以把它总结成如下的定理．

定理 I 设线性关系

$$\mathbf{x}=C\mathbf{y}$$

定义直角坐标系的一个变换，则 C 的列在旧坐标系(即 x_i 坐标系)中表示新轴方向的单位向量．

该定理在二维与三维情况下总是正确的．此时，根据维数，C

是 2×2 矩阵或 3×3 矩阵.

定理 I 的推论 设 $C=(c_{ij})$ 在三维空间中定义直角坐标系的一个变换,则

$$c_{1j}^2+c_{2j}^2+c_{3j}^2=1 \quad (j=1,2,3), \tag{6}$$

$$c_{1j}c_{1k}+c_{2j}c_{2k}+c_{3j}c_{3k}=0 \quad (\text{当 } j\neq k \text{ 时}). \tag{7}$$

二维情况下的对应公式可从上面各式中去掉元素 c_{3j} 与 c_{i3} 而得到.

证明 关系式(6)与(7)只不过分别表示沿新轴方向的单位向量长度为 1 与两两正交这两个事实.

定义 I 如果一个 3×3 (或 2×2) 矩阵 C 的列看成为在一个直角坐标系中表示出来的向量时,是两两正交的单位向量,也就是说矩阵 C 满足定理 I 的推论中的条件,则称 C 为正交矩阵.

因此,一个矩阵若定义了直角坐标系的变换,则它是正交矩阵.

可以用一个简洁的方法把(6)、(7)两式用矩阵形式表达出来.考虑 C 与其转置 C' 的乘积

$$C'C=\begin{pmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{pmatrix}\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

若把它们乘出来并且与(6)、(7) 式相比,我们看到,C 是正交矩阵当且仅当

$$C'C=I. \tag{8}$$

所以我们有如下的定理.

定理 II 一个 3×3(或 2×2) 实矩阵 C 是正交矩阵当且仅当该矩阵是非奇异的,并且它的逆矩阵等于它的转置矩阵:

$$C^{-1}=C'. \tag{9}$$

例 2 下列矩阵中哪些是正交矩阵?

$$\text{(i)}\begin{pmatrix}\frac{3}{5} & \frac{4}{5}\\ \frac{4}{5} & -\frac{3}{5}\end{pmatrix};\quad \text{(ii)}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix};$$

$$\text{(iii)}\begin{pmatrix}1 & 2 & -2\\ 2 & 1 & 2\\ -2 & 2 & 1\end{pmatrix}.$$

矩阵(i)与(ii)是正交矩阵．矩阵(iii)不是正交矩阵；虽然它的各列互相正交，但每个列向量的长度是3．因此，若把矩阵(iii)的每个元素都乘以$\frac{1}{3}$，所得到的矩阵就是正交矩阵了．

注意，若在(ii)中置$\alpha=45°$，就得到本节例1中的矩阵(5)．

早些时候我们曾见过矩阵(ii)，不过在一个不同的场合——在一个固定的直角坐标系中，它是一个旋转的矩阵（见 9.3 节）．一个矩阵C在几何中能够扮演这样两种角色，把它们区别清楚是很重要的：

(i) 矩阵C可以表示在一个给定的坐标系中的线性变换（第九章）．此时，我们采用记号$\mathbf{x}'=C\mathbf{x}$，其中$\mathbf{x}'$表示变换后的向量．坐标系自始至终固定不变．

(ii) 矩阵C可以定义一个坐标变换（本节）．此时，我们采用记号$\mathbf{x}=C\mathbf{y}$（或等价地用$\mathbf{y}=C^{-1}\mathbf{x}$），这里，一个固定的向量在旧坐标系与新坐标系中分别用$\mathbf{x}$与$\mathbf{y}$表示．

练　习

1．下列矩阵中哪些是正交矩阵？

$$\text{(i)}\ \frac{1}{13}\begin{pmatrix}5 & 12\\ -12 & 5\end{pmatrix};\quad \text{(ii)}\begin{pmatrix}1 & 1 & -1\\ 0 & 2 & 1\\ 1 & -1 & 1\end{pmatrix};\quad \text{(iii)}\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & -1\\ -1 & 0 & 0\end{pmatrix};$$

(iv) 恒等矩阵．

2．写出练习1中各正交矩阵的逆矩阵。

3．求矩阵

$$A=\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$$

的逆矩阵。

$\left[\text{提示：注意及}\frac{1}{3}A\text{ 是正交矩阵,用公式(9)}.\right]$

4．证明，若 C 是正交矩阵，则它的转置 C' 也是正交矩阵。

5．证明，若一个矩阵的列是两两正交的单位向量，则它的行也具有同样的性质。

6．在直角坐标系 Ox_1，Ox_2 中求以 O 为中心的一个椭圆的方程，已知该椭圆的长轴与 Ox_1 轴倾成 α 角，这里 $\tan\alpha=\frac{3}{4}$，并且长、短二轴的长度分别为 $2a$ 与 $2b$。

14.2 线性变换与坐标变换的矩阵表示

设 T 是一个线性变换（例如射影、旋转或反射），在给定的坐标系中由矩阵 A 表示．此时，一个向量 $\mathbf{x}$ 及其在 T 下的像 $\mathbf{x}'$ 的关系是

$$\mathbf{x}'=A\mathbf{x}. \tag{10}$$

假如我们变换坐标系，使得在旧坐标系中用 $\mathbf{x}$ 表示的向量变成 $\mathbf{y}$，其中

$$\mathbf{x}=C\mathbf{y}. \tag{11}$$

把(11)式代入(10)式，我们得到

$$C\mathbf{y}'=AC\mathbf{y},$$

于是

$$\mathbf{y}'=C^{-1}AC\mathbf{y}. \tag{12}$$

关系式(12)刻划了在新坐标系中 T 的作用情况．我们已经证明了如下的定理：

定理 III 设T为线性变换，在"旧"坐标系中与在"新"坐标系中分别用矩阵A与B表示，则

$$B=C^{-1}AC,$$

其中C是（按关系式(11)）定义坐标变换的矩阵．

注意A与B是相似矩阵．

例 3

我们已经知道（参看 9.3 节的例 2），到与水平轴倾成 45°角的直线L上的正交射影（图 83）在x_1, x_2坐标系中是由矩阵

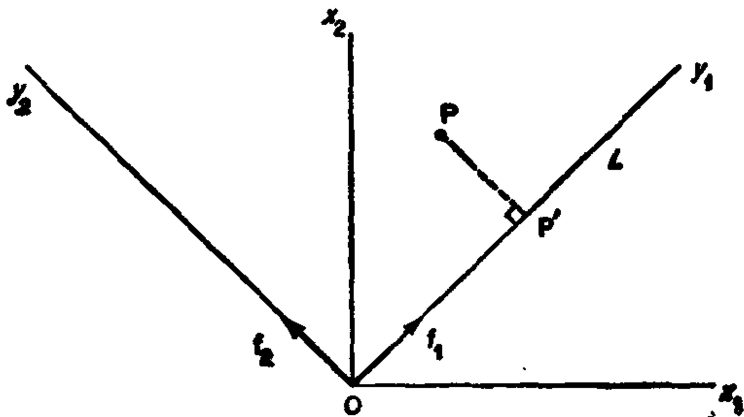


图 83 到L上的正交射影

$$A=\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}$$

给出的．现在把坐标轴旋转 45°，得到y_1，y_2坐标系．在旧坐标系中用$\mathbf{x}$表示的向量在新坐标系中则用$\mathbf{y}$表示，这里

$$\mathbf{x}=C\mathbf{y},\quad C=\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}$$

（应用 14.1 节的定理 I）．

在新坐标系中，这个正交射影由矩阵

$$C^{-1}AC=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$$

给出．这个矩阵形式之简单是不足为奇的．它的列只不过是与如下事实相吻合的：这个射影把单位向量 $\mathbf{f}_1$ 与 $\mathbf{f}_2$ 分别映到 $\mathbf{f}_1$ 与 $\mathbf{0}$．

下面这个例题引导我们得出一个简单的公式，以求由一个旋转矩阵所定义的旋转的角度．

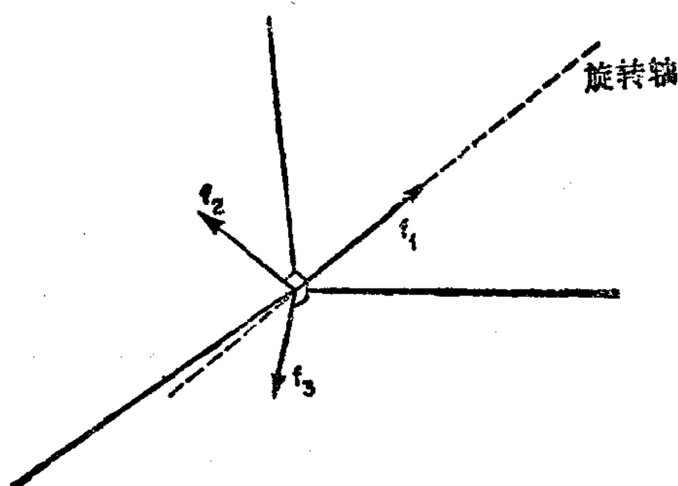


图 84 旋转问题中的坐标系

例 4

再考虑矩阵(9.6 节的例 9)

$$R=\begin{pmatrix}\frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{2}}{2} & -\frac{1}{2} & 0\\ \frac{\sqrt{3}}{4} & \frac{3}{4} & -\frac{1}{2}\end{pmatrix},$$

在一个直角坐标系中，它按关系式

$$\mathbf{x}'=R\mathbf{x} \tag{13}$$

定义了一个旋转．现在引进一个新的直角坐标系，使得旋转轴为新的坐标轴之一．具体地说，假设新坐标轴方向的单位向量为 $\mathbf{f}_1$，$\mathbf{f}_2$ 与 $\mathbf{f}_3$，其中 $\mathbf{f}_1$ 在旋转轴上．在本例中，$\mathbf{f}_1$ 在旧坐标系中应表为

$$\mathbf{f}_1=\frac{1}{\sqrt{5}}\begin{pmatrix}\sqrt{3}\\1\\1\end{pmatrix}.$$

这个旋转

(i) 保持 $\mathbf{f}_1$ 不动；

(ii) 对 $\mathbf{f}_2$ 与 $\mathbf{f}_3$ 的作用就象一个二维空间的旋转．

因此，对于基底 $\mathbf{f}_1, \mathbf{f}_2$ 与 $\mathbf{f}_3$，该旋转的矩阵是

$$B=\begin{pmatrix}1 & 0 & 0\\0 & \cos\alpha & -\sin\alpha\\0 & \sin\alpha & \cos\alpha\end{pmatrix},$$

其中 α 是旋转角．根据定理 III，

$$B=C^{-1}RC,$$

其中 C 是坐标变换矩阵．因为 B 与 R 相似，它们的迹相同(根据 13.3 节，定理 VII 的推论)．但

$$B\text{ 的迹}=1+\cos\alpha+\cos\alpha=1+2\cos\alpha,$$

因此

$$1+2\cos\alpha=R\text{ 的迹}=\frac{1}{4}-\frac{1}{2}-\frac{1}{2}=-\frac{3}{4},$$

所以

$$\cos\alpha=-\frac{7}{8}.$$

这又验证了 9.6 节的(19)式．

例 4 中的论证过程是相当一般化的，因此，我们得到如下的结果．

定理 IV 设 R 是 3×3 的旋转矩阵，则其旋转角 α 由

$$1+2\cos\alpha = R \text{ 的迹}$$

给出.

练　　习

7. 已知矩阵

$$\frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix}$$

表示一个旋转，求其旋转轴与旋转角.

8. 证明，旋转矩阵的行列式值为 1.

9. 证明，若 3×3 矩阵 S 表示一个反射，则存在一个正交矩阵 C，使

$$C^{-1}SC=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

当

$$S=\frac{1}{9}\begin{pmatrix} 7 & 4 & -1 \\ 4 & 1 & 8 \\ -4 & 8 & 1 \end{pmatrix}$$

时，求这样的矩阵 C.

10. 证明，反射矩阵的行列式值为 -1. 它的迹是多少？

11. 设 3×3 矩阵 A 满足条件 $A'A=I$. 证明，有如下的向量距离关系式：

$$\|\mathbf{x}-\mathbf{y}\|=\|A\mathbf{x}-A\mathbf{y}\|.$$

[提示：用 $\|\mathbf{v}\|^2=\mathbf{v}'\mathbf{v}$.]

由此推出，一个 3×3 正交矩阵可以看成表示一个保持原点不动的等距变换的矩阵，亦即，它表示一个旋转或一个反射.

对 2×2 矩阵，类似结果也成立.

14.3　正交群

定理 V　所有 2×2(或 3×3)正交矩阵的集合在矩阵乘法下成为一个群，称为二维(或三维)正交群.

证明 我们来验证封闭性. 设 C_1 与 C_2 是正交矩阵,则

$$C_1'C_1=I, \quad C_2'C_2=I.$$

此时

$$(C_1C_2)'(C_1C_2)=C_2'C_1'C_1C_2=C_2'IC_2=I,$$

所以 C_1C_2 也是正交矩阵.

恒等矩阵 I 是正交矩阵.

剩下的是要验证逆元素公理. 如果 C 是正交矩阵, 则 $C'=C^{-1}$,取其转置,$C=(C^{-1})'$. 这意味着

$$(C^{-1})'=(C^{-1})^{-1},$$

因此, C^{-1} 也是正交矩阵. 证毕.

我们已经知道(14.2 节的练习 11), 一个二维或三维的正交矩阵表示一个旋转或一个反射,因此,二维正交群由所有形如

$$\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\ (\text{转 }\alpha\text{ 角的旋转})$$

与

$$\begin{pmatrix}\cos\beta & \sin\beta\\ \sin\beta & -\cos\beta\end{pmatrix}\ \left(\text{关于 } x_2=\tan\left(\frac{1}{2}\beta\right)x_1 \text{ 的反射}\right)$$

矩阵组成, 这里 α 与 β 取遍所有实数. 注意旋转矩阵的行列式值为 $+1$, 而反射矩阵的行列式值为 -1.

现在考虑三维正交群. 根据 14.2 节例 4 的论证,每一个 3×3 旋转矩阵 R 都相似于一个形如

$$B=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha\\ 0 & \sin\alpha & \cos\alpha\end{pmatrix}$$

的矩阵. 因为相似矩阵的行列式值相同, 所以

$$|R|=|B|=1.$$

另一方面, 每一个 3×3 反射矩阵 S 都相似于

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(证明：引进新的直角坐标系 $\mathbf{f}_1$, $\mathbf{f}_2$ 与 $\mathbf{f}_3$，使 $\mathbf{f}_1$ 垂直于反射平面，而 $\mathbf{f}_2$ 与 $\mathbf{f}_3$ 都在反射平面内.）所以 $|S|=-1$. 我们证明了：

定理 VI 正交群中的每个矩阵依据它表示旋转还是反射而分别具有行列式值 $+1$ 或 -1.

练 习

12. 证明，正交群中行列式值为 1 的矩阵组成一个子群. 行列式值为 -1 的矩阵也是这样吗？

13. 证明，三维正交群中，如果一个矩阵的行列式值为 -1，则它的迹为 1. 在这个群中还有哪些矩阵的迹是 1？

14.4 正交矩阵的推广

正交矩阵的定义可以毫无困难地推广到 $n\times n$ 实矩阵或复矩阵.

定义 IIa 如果一个 $n\times n$ 实矩阵 C 的列看成欧几里得空间 $\mathscr{E}^n$ 的向量时，是两两正交的单位向量，也就是说，它们组成 $\mathscr{E}^n$ 的一组正规正交基，则称 C 为正交矩阵.

定义 IIb 如果一个 $n\times n$ 复矩阵 U 的列看成酉空间 $\mathscr{U}^n$ 的向量时，它们组成 $\mathscr{U}^n$ 的一组正规正交基，则称 U 为酉矩阵.

对实矩阵，仍保留"正交"这一术语. 实的酉矩阵是正交矩阵. 因此，对酉矩阵已经证明的任何结果，对(实)正交矩阵都自动成立.

一个 $n\times n$ 实矩阵 $C=(c_{ij})$ 是正交矩阵当且仅当

$$c_{1j}c_{1k}+c_{2j}c_{2k}+\cdots+c_{nj}c_{nk}=\begin{cases}1, & \text{当 } j=k \text{ 时},\\ 0, & \text{当 } j\neq k \text{ 时}.\end{cases} \tag{15}$$

一个 $n\times n$ 复矩阵 $U=(u_{ij})$是酉矩阵当且仅当

$$\bar{u}_{1j}u_{1k}+\bar{u}_{2j}u_{2k}+\cdots+\bar{u}_{nj}u_{nk}=\begin{cases}1, & \text{当 } j=k \text{ 时},\\ 0, & \text{当 } j\neq k \text{ 时}.\end{cases} \tag{16}$$

(参看 5.6 节的 (25) 式.) 下面的定理是这些关系式的直接推论.

定理 VII 一个矩阵U是酉矩阵当且仅当它是非奇异的，且它的逆矩阵等于它的共轭转置：

$$U^{-1}=\bar{U}'. \tag{17}$$

证明 容易看出，关系式(16)当且仅当

$$\bar{U}'U=I \tag{18}$$

时满足，由此即得(17)式.

例 5 下列矩阵中哪些是酉矩阵？哪些是正交矩阵？

$$\text{(i)}\ \frac{1}{2}\begin{pmatrix}1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\\ 1 & 1 & -1 & -1\\ 1 & 1 & 1 & 1\end{pmatrix};\quad \text{(ii)}\ \begin{pmatrix}1 & 1\\ i & i\end{pmatrix};$$

$$\text{(iii)}\begin{pmatrix}\frac{1+i}{2} & \frac{1-i}{2} & 0\\ \frac{i}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}}\\ \frac{1}{2} & \frac{i}{2} & \frac{i}{\sqrt{2}}\end{pmatrix}.$$

(i)与(iii)是酉矩阵. 要验证(iii)，我们先计算其列向量 $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ 的长度

$$\|\mathbf{u}_1\|^2=\overline{\frac{1+i}{2}}\,\frac{1+i}{2}+\overline{\frac{i}{2}}\,\frac{i}{2}+\overline{\frac{1}{2}}\,\frac{1}{2}$$

$$=\frac{1}{4}\{(1-i)(1+i)+(-i)i+1\}=1.$$

因此，$\|\mathbf{u}_1\|=1$．同理可得$\|\mathbf{u}_2\|=\|\mathbf{u}_3\|=1$．

接着，我们验证列向量之间的正交性

$$\langle\mathbf{u}_1,\ \mathbf{u}_2\rangle=\overline{\frac{1+i}{2}}\,\frac{1-i}{2}+\overline{\frac{i}{2}}\,\frac{-1}{2}+\overline{\frac{1}{2}}\,\frac{i}{2}$$

$$=\frac{1}{4}\{(1-i)(1-i)+i+i\}=0.$$

因此，$\mathbf{u}_1$ 与 $\mathbf{u}_2$ 正交．其余的正交性同理可证．

矩阵(ii)不是酉矩阵，但是，若把它的每个元素都乘以$\frac{1}{\sqrt{2}}$，得到的矩阵是酉矩阵．

矩阵(i)是正交矩阵，因为它是实的酉矩阵．其余矩阵都不能称为正交矩阵．

在本节结束之前，我们再简单地讨论一下酉矩阵的某些性质．

定理 VIIIa 酉矩阵的行列式的绝对值为 1．

证明 这个证明依赖于上文的两个定理：矩阵乘积的行列式的定理与转置矩阵的行列式的定理．设U是酉矩阵，则$\overline{U}'U=I$．取行列式，我们有

$$|I|=1=|\overline{U}'U|=|\overline{U}'|\,|U|=\overline{|U'|}\,|U|=\overline{|U|}\,|U|.$$

因此，$|U|$的绝对值为 1．

对正交矩阵C，上述的计算给出

$$|C|^2=1.$$

定理 VIIIb 正交矩阵的行列式值为 +1 或 −1．

定理 IX

(a) 酉矩阵所有特征值的绝对值都是 1．

(b) 对应不同特征值的特征向量互相正交．

这个定理的证明需要如下的引理.

引理 设U是一个酉矩阵,则对于所有 $\mathbf{x},\mathbf{y}$,有

$$\langle U\mathbf{x},U\mathbf{y}\rangle=\langle\mathbf{x},\mathbf{y}\rangle. \tag{19}$$

引理的证明 据 10.4 节的关系式(10.19),有

$$\langle U\mathbf{x},U\mathbf{y}\rangle=\langle\overline{U}'U\mathbf{x},\mathbf{y}\rangle.$$

于是,从关系式(18)即得引理.

定理 IX 的证明 (a)设 $\mathbf{v}$ 是酉矩阵U的特征向量,对应特征值λ. 用(19)式,我们得到

$$\langle\mathbf{v},\mathbf{v}\rangle=\langle U\mathbf{v},U\mathbf{v}\rangle=\langle\lambda\mathbf{v},\lambda\mathbf{v}\rangle=\overline{\lambda}\lambda\langle\mathbf{v},\mathbf{v}\rangle.$$

因为$\langle\mathbf{v},\mathbf{v}\rangle=\|\mathbf{v}\|^2\neq0$,所以 $\overline{\lambda}\lambda=|\lambda|^2=1$,正如所求.

(b) 假定 $\mathbf{v}_1$, $\mathbf{v}_2$ 是U的特征向量,分别对应不同的特征值λ_1,λ_2. 由(19)式

$$\langle\mathbf{v}_1,\mathbf{v}_2\rangle=\langle U\mathbf{v}_1,U\mathbf{v}_2\rangle=\langle\lambda_1\mathbf{v}_1,\lambda_2\mathbf{v}_2\rangle=\overline{\lambda}_1\lambda_2\langle\mathbf{v}_1,\mathbf{v}_2\rangle. \tag{20}$$

把(20)式两边同乘以 λ_1, 忆及 $\overline{\lambda}_1\lambda_1=1$,我们得到

$$\lambda_1\langle\mathbf{v}_1,\mathbf{v}_2\rangle=\lambda_2\langle\mathbf{v}_1,\mathbf{v}_2\rangle.$$

因为 $\lambda_1\neq\lambda_2$, 所以$\langle\mathbf{v}_1,\mathbf{v}_2\rangle=0$. 证毕.

例 6 求酉矩阵

$$U=\frac{1}{2}\begin{pmatrix}i-1 & 1-i\\ i-1 & i-1\end{pmatrix}$$

的特征值和对应的特征向量.

U的特征多项式是

$$|U-\lambda I|=\lambda^2-(i-1)\lambda-i=(\lambda+1)(\lambda-i).$$

所以U的特征值是-1与i. 可以验证,$\mathscr{S}_{-1}$与$\mathscr{S}_i$分别由向量$\begin{pmatrix}1\\ -i\end{pmatrix}$与$\begin{pmatrix}1\\ i\end{pmatrix}$生成.

正如定理所述,U的特征值的绝对值都是 1,且 $\mathscr{S}_{-1}$ 中的向量与 $\mathscr{S}_i$ 中的向量正交.

练　习

14．下列矩阵中哪些是酉矩阵？

(i) $\begin{pmatrix} i-1 & 1-i \\ i-1 & i-1 \end{pmatrix}$；(ii) $\frac{1}{2}\begin{pmatrix} i-1 & 1-i \\ i-1 & i-1 \end{pmatrix}$；(iii) $\begin{pmatrix} \frac{1+i}{2} & \frac{i}{2} & \frac{1}{2} \\ \frac{1-i}{2} & -\frac{1}{2} & \frac{i}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}$；

(iv) 任意一个 $n\times n$ 置换矩阵。

15．验证共轭矩阵的下述性质。

(i) $\overline{\overline{A}}=A$；

(ii) $\overline{AB}=\overline{A}\,\overline{B}$；

(iii) $\overline{A^{-1}}=(\overline{A})^{-1}$；

[提示：$\overline{AA^{-1}}=\overline{I}=I$，再用性质(ii)。]

(iv) $(\overline{A})'=\overline{A'}$；

(v) $|\overline{A}|=\overline{|A|}$。

在定理 VIII 的证明中曾用过性质(iv)与(v)。

16．证明，如果U是酉矩阵，则 $\overline{U}, U', -U$ 与 U^{-1} 也都是酉矩阵。

17．求酉矩阵

$$\begin{pmatrix} \frac{1}{2}(1+i) & -\frac{1}{2}(1+i) \\ \frac{1}{2}(1+i) & \frac{1}{2}(1+i) \end{pmatrix}$$

的特征值，证实它们的绝对值都是 1，并且证明，对应不同特征值的特征向量(看成$\mathscr{U}^2$的向量)互相正交。

18．对正交矩阵

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

重复练习 17。

第十五章　二次型与对称矩阵

15.1　引论

对称矩阵的重要性在于它们与二次型的联系(参看 10.4 节)。二次型有许多应用,我们先举几个例子,作为本章的开始.

例 1　动能与势能

我们来计算 13.1 节例 3 中的振动弹簧的动能与势能. 用两个质点偏离平衡位置的位移 x_1 与 x_2 来表示这一系统的动能与势能,我们有

$$\text{系统的动能} = T = \frac{1}{2}m(\dot{x}_1^2+\dot{x}_2^2),$$

$$\text{系统的势能}^* = V = \frac{1}{2}kx_1^2+\frac{1}{2}k(x_2-x_1)^2+\frac{1}{2}k(-x_2)^2$$

$$= k(x_1^2+x_2^2-x_1x_2).$$

动能是 $\dot{x}_1$ 与 $\dot{x}_2$ 的二次型

$$T=\dot{\mathbf{x}}^t K\dot{\mathbf{x}},\quad \text{其中 } K=\begin{pmatrix}\frac{1}{2}m & 0\\ 0 & \frac{1}{2}m\end{pmatrix}.$$

势能是 x_1 与 x_2 的二次型

---

* 把弹性系数为 k 的弹簧从其自然长度拉长 s,所做的功是

$$\int_0^s kx\,dx=\frac{1}{2}ks^2$$

把这个公式分别用于三段弹簧。

$$V=\mathbf{x}'A\mathbf{x},\ 其中\ A=\begin{pmatrix} k & -\frac{1}{2}k \\ -\frac{1}{2}k & k \end{pmatrix}.$$

该系统完全被矩阵 K 与 A 所刻划. 因此，我们希望正规振动方式的角频率可以从这两个矩阵得出. 事实上，可以证明，角频率是 $\omega=\sqrt{\mu}$，其中 μ 是行列式方程

$$|A-\mu K|=0$$

的解. 进一步的讨论，参看 16.4 节.

例 2 圆锥曲线

二次方程

$$a_{11}x_1^2+a_{22}x_2^2+2a_{12}x_1x_2+b_1x_1+b_2x_2+c=0 \qquad (1)$$

表示圆锥曲线，在非退化的情况下*，它是椭圆、双曲线或抛物线. 反过来，任一圆锥曲线的方程都可写成(1)的形式. 圆锥曲线(1)的性质是由二次型

$$a_{11}x_1^2+a_{22}x_2^2+2a_{12}x_1x_2=\mathbf{x}'A\mathbf{x}$$

决定的. 要判断方程(1)表示的是椭圆、双曲线还是抛物线，只要求出对称矩阵 A 的行列式值就够了(参看 16.2 节).

例 3 空间的度量

三维空间中弧长元素 ds 是由公式

$$ds^2=\begin{cases} dx^2+dy^2+dz^2 & (在直角坐标系中), \\ d\rho^2+\rho^2d\theta^2+dz^2 & (在柱面坐标系中), \\ dr^2+r^2\sin^2\phi d\theta^2+r^2d\phi^2 & (在球极坐标系中) \end{cases}$$

给出的. ds^2 的表达式称为该空间的度量，该表达式与所使用的坐标系有关.

---

* 方程(1)有时可能分解成两个线性因子，在这种情况下，它表示两条直线；此外，它还可能不表示任何轨迹，例如

$$x_1^2+x_2^2+1=0.$$

与狭义相对论相适应的度量是

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2, \tag{2}$$

这里 x, y, z 是在一个直角坐标系中测量得到的长度，t 表示时间，而 c 是光速常数.

上述所有度量都是二次型，并且是特别简单的二次型，因为这些二次型的对称矩阵都是对角矩阵. 更一般地说，一个空间(以四维空间为例)的度量是由如下形式的一个二次型所定义的：

$$ds^2 = \sum_{i=1}^{4} \sum_{j=1}^{4} g_{ij} dx_i dx_j.$$

例 4　刚体力学：惯性矩阵

考虑一个运动的刚体，设在运动过程中刚体上有一点（记作 O）始终保持不动. 这个刚体的动能也是一个二次型，并且用矩阵工具不难把它求出. 我们现在就来说明这一点.

刚体上一点 P 的速度 $\mathbf{v}$ 由

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$$

给出(4.4 节的例 15)，这里 $\mathbf{r}$ 是点 P 关于点 O 的位置向量，而 $\boldsymbol{\omega}$ 是在通过 O 的瞬时旋转轴方向上的角速度向量.

引进一个以 O 为原点的直角坐标系，我们有

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \boldsymbol{\omega} = \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}, \quad \text{以及 } \boldsymbol{\omega} \times \mathbf{r} = \begin{pmatrix} \omega_y z - \omega_z y \\ \omega_z x - \omega_x z \\ \omega_x y - \omega_y x \end{pmatrix}. \tag{3}$$

最后一个表达式可以写成矩阵形式

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} = S\boldsymbol{\omega}, \quad \text{这里 } S = \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}. \tag{4}$$

由此可见，点 P 的速率由公式

$$\|\mathbf{v}\|^2 = \mathbf{v}'\mathbf{v} = \boldsymbol{\omega}' S' S \boldsymbol{\omega} \tag{5}$$

给出，这里

$$S'S=\begin{pmatrix}0 & -z & y\\ z & 0 & -x\\ -y & x & 0\end{pmatrix}\begin{pmatrix}0 & z & -y\\ -z & 0 & x\\ y & -x & 0\end{pmatrix}$$

$$=\begin{pmatrix}y^2+z^2 & -xy & -xz\\ -yx & x^2+z^2 & -yz\\ -zx & -zy & x^2+y^2\end{pmatrix}. \qquad (6)$$

假定这一刚体有连续的密度分布 $\rho=\rho(x,y,z)$，则它的动能是

$$T=\iiint\frac{1}{2}\rho\|\mathbf{v}\|^2dxdydz, \qquad (7)$$

积分是对所考虑的瞬间该物体所占据的空间区域取的。

把(5)、(6)两式代入(7)，我们得到二次型

$$T=\frac{1}{2}\omega' J\omega, \qquad (8)$$

其中

$$J=\begin{pmatrix}A & -H & -G\\ -H & B & -F\\ -G & -F & C\end{pmatrix}, \qquad (9)$$

而

$$A=\iiint(y^2+z^2)\rho dxdydz,$$

$$B=\iiint(x^2+z^2)\rho dxdydz,$$

$$C=\iiint(x^2+y^2)\rho dxdydz,$$

$$F=\iiint yz\rho dxdydz,$$

$$G=\iiint xz\rho dxdydz,$$

$$H=\iiint xy\rho dxdydz.$$

积分式 A, B 与 C 分别称为该刚体关于 Ox 轴, Oy 轴与 Oz 轴的惯性矩*, 而 F, G, H 称为惯性积. 矩阵 J 称为该刚体在所给的坐标系中的惯性矩阵.

如果把坐标系固定于空间，惯性矩阵就与刚体的运动有关. 为了避免复杂化，通常采用固定在刚体上的坐标轴组成的参照系. 这样得到的运动方程称为欧拉方程.

注意，若旋转轴与 Ox 轴重合，在这一特殊情况下有

$$\omega=\begin{pmatrix}\omega\\0\\0\end{pmatrix},$$

而动能则由众所周知的公式

$$T=\frac{1}{2}I\omega^2$$

给出，其中 I 等于该刚体关于它的旋转轴的惯性矩 A.

在停止这个例子的讨论之前，我们还想作几点评注.

(i) (4)式所定义的矩阵满足关系式

$$S'=-S, \tag{10}$$

所以它是斜称矩阵.

(ii) 惯性矩阵 J 是对称矩阵. 从(6)式显然可以看出, J 可以表为一个矩阵 M 与它的转置矩阵的乘积 $M'M$. 所有像这样的乘积都是对称矩阵.

(iii) 对于本题所述的刚体，其角动量有一个表达式，使得惯性矩阵在其中起重要作用. 由 4.4 节，我们有

$$角动量=\iiint \mathbf{r}\times\mathbf{v}\rho\,dx\,dy\,dz. \tag{11}$$

而

* 惯性矩(moment of inertia)又译作转动惯量——译者注.

$$\mathbf{r}\times\mathbf{v}=-\mathbf{v}\times\mathbf{r}=(-S\boldsymbol{\omega})\times\mathbf{r}=S(-S\boldsymbol{\omega}).$$

其中最后一步用了(4)式，但把(4)式中的向量 $\boldsymbol{\omega}$ 换成 $-S\boldsymbol{\omega}$。从(10)式得知 $-S=S'$，因此 $-S^2=S^tS$。所以

$$\mathbf{r}\times\mathbf{v}=S^tS\boldsymbol{\omega}.$$

以此式代入公式(11)，我们得到

$$角动量=J\boldsymbol{\omega}.$$

例 5 设有一个长方体形状的刚性框架，其本身质量可以忽略不计.在每个顶点都放上质量 m. 又设长方体各棱长分别为 a, b, c. 取一个坐标系，使该长方体有三条棱都在坐标轴上(图 85). 在此坐标系中，求该刚体的惯性矩阵.

为了辨认起见，我们把各顶点编号，并且把顶点 i 的坐标记为 (x_i,y_i,z_i)，在此顶点上的质量记为 m_i。当然，对所有的 i 均有 $m_i=m$。

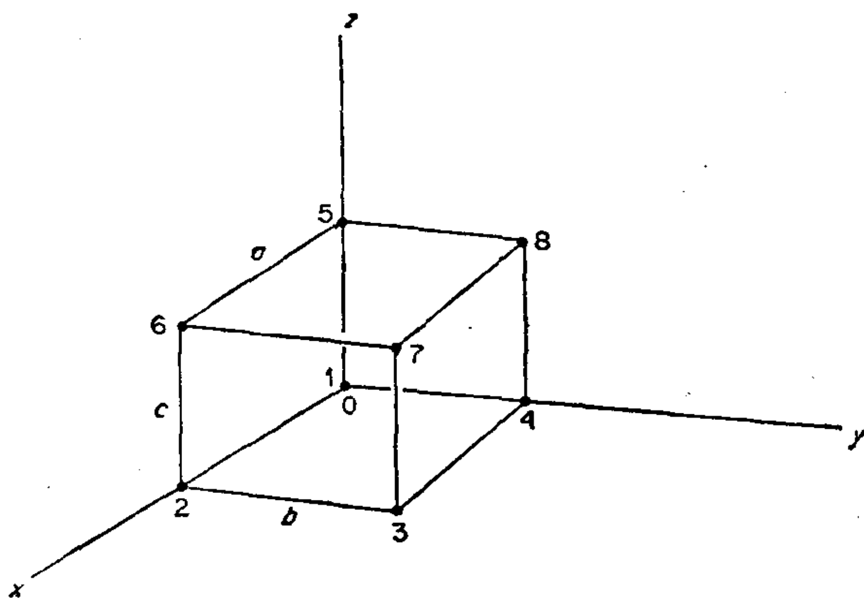


图 85 顶点已编号的长方体

本例中，惯性矩与惯性积的公式里的积分换成了求和. 我们有

$$A=\sum_{i=1}^{8} m_i(y_i^2+z_i^2)$$

$$=(m_3+m_4)(b^2+0^2)+(m_5+m_6)(0^2+c^2)+(m_7+m_8)(b^2+c^2)$$

$$=4\,m(b^2+c^2).$$

又

$$H=\sum_{i=1}^{8} m_i x_i y_i$$

$$=(m_3+m_7)ab=2\,mab.$$

照此方法继续计算，我们得到惯性矩阵

$$J=\begin{pmatrix} 4\,m(b^2+c^2) & -2\,mab & -2\,mac \\ -2\,mab & 4\,m(a^2+c^2) & -2\,mbc \\ -2\,mac & -2\,mbc & 4\,m(a^2+b^2) \end{pmatrix}.$$

练　　习

1．把一条弹性系数为 k 的弹簧放在光滑的水平桌面上，让其处于自然状态，并固定它的两端．再把质量 m 与 $\frac{1}{2}m$ 的两个质点分别加在该弹簧的两个三等分点上．如果该系统在弹簧所在的直线上振动，求该系统的动能与势能，请用两个质点偏离平衡位置的位移 x_1 与 x_2 表出．

2．若以质量 m，$2\,m$ 与 m 的三个质点分别加在练习 1 所说的弹簧的三个四等分点上，重做练习 1．

3．设有一个质量可以忽略不计的单位正立方体，在其各顶点都放上相同的质量 m．取一个坐标系，使 Ox，Oy，Oz 三轴上各有一条棱．在此坐标系中求该正方体的惯性矩阵．

4．设例 3 所说的正方体以 ω 弧度/秒的角速度

(i) 绕 Oz 轴旋转；

(ii) 绕通过 O 点与 P (1,1,1) 点的直线旋转；

(iii) 绕通过 O 点与 Q (1,1,0)点的直线旋转．

分别求该正方体的动能与角动量．

15.2 实对称矩阵的特征值

在 13.3 节中，我们曾研究了把一个矩阵对角化的问题，并且以解一个动力学系统的问题为例，指出了它的重要性。我们也看到了还存在一些不相似于对角矩阵的正方矩阵。我们现在要证明，实对称矩阵总是可对角化的。事实上，我们将要证明的还稍微更多一些。

设 A 是实对称矩阵。我们要证明：

(i) A 的特征值全是实数；

(ii) 矩阵 A 相似于一个对角矩阵；更进一步说，存在一个正交矩阵 C，使 $C^{-1}AC$ 是对角矩阵。

我们将要看到，这些结果在研究二次型时是很重要的。

定理 Ia 实对称矩阵 A 的所有特征值都是实数。

证明 设 λ 是 A 的一个特征值，$\mathbf{v}$ 是对应于 λ 的特征向量，则

$$A\mathbf{v}=\lambda\mathbf{v} \quad (\mathbf{v}\neq\mathbf{0}).$$

可以设想 λ 是一个复数而 $\mathbf{v}$ 的某些分量是复的。我们先承认这种可能性，即假定在酉空间 $\mathscr{U}^n$（假定 A 是 $n\times n$ 的）中讨论。因为 A 是实对称矩阵，所以它是埃尔米特矩阵，从而 $\langle\mathbf{v},A\mathbf{v}\rangle$ 是实的（参看 10.4 节的例 5）。此时

$$\langle\mathbf{v},A\mathbf{v}\rangle=\langle\mathbf{v},\lambda\mathbf{v}\rangle=\lambda\langle\mathbf{v},\mathbf{v}\rangle,$$

所以

$$\lambda=\frac{\langle\mathbf{v},A\mathbf{v}\rangle}{\|\mathbf{v}\|^2},$$

这是一个实数。证毕。

上面的证明无须变动就可适用于所有埃尔米特矩阵。所以我们有

定理 Ib 埃尔米特矩阵的所有特征值都是实数.

定理 I 的推论 实对称矩阵,或更一般地说,埃尔米特矩阵的特征多项式可分解为实线性因子的积.

证明 否则该矩阵有复特征值.

例 6

把定理 I 同 13.2 节各例题对照. 例 5 与例 6 的矩阵是实对称矩阵,它们的特征值都是实的. 另一方面,旋转矩阵(例 8)有复特征值,当然它不是对称矩阵. 注意,也存在一些实的非对称矩阵,它的所有特征值都是实的,切变矩阵(例 7)就是一例.

例 7 求下列埃尔米特矩阵的特征值:

$$\text{(i)}\ \begin{pmatrix} 1 & i \\ -i & -2 \end{pmatrix};\qquad \text{(ii)}\ \begin{pmatrix} 5 & 4-3i & -2-6i \\ 4+3i & 5 & 2-6i \\ -2+6i & 2+6i & 8 \end{pmatrix}.$$

矩阵(i)的特征方程为

$$\begin{vmatrix} 1-\lambda & i \\ -i & -2-\lambda \end{vmatrix} = \lambda^2+\lambda-3=0,$$

所以它的特征值为$\frac{1}{2}(-1\pm\sqrt{13})$.

可以求得矩阵(ii)的特征方程为 $\lambda^2(\lambda-18)=0$. 所以它的特征值是 0 与 18.

在 13.2 节的定理 IV 中,我们曾证明,对任一矩阵 A,对应不同特征值的特征向量线性无关. 当 A 是实对称矩阵或埃尔米特矩阵时,我们可以说得更多一些:

定理 II 实对称矩阵(或埃尔米特矩阵)A 的对应不同特征值 $\lambda_1,\cdots,\lambda_r$ 的特征向量 $\mathbf{v}_1,\cdots,\mathbf{v}_r$ 组成一个正交集.

证明 我们必须证明,对任意 $i,j(i\neq j)$,$\mathbf{v}_i$ 与 $\mathbf{v}_j$ 正交. 根据 10.4 节的关系式(19)

$$\langle\mathbf{v}_i, A\mathbf{v}_j\rangle=\langle A\mathbf{v}_i, \mathbf{v}_j\rangle.$$

因此

$$\langle \mathbf{v}_i, \lambda_j \mathbf{v}_j \rangle = \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle,$$

可见

$$\lambda_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \bar{\lambda}_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \lambda_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle,$$

最后一步是由于 λ_i 为实数．因为 $\lambda_i \neq \lambda_j$，由此得出$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$，这就是所要证的．

例 8

我们以 13.2 节的例 5 与例 6 的对称矩阵为例来说明定理 II．

在例 5 中，$A=\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$的特征值是$-1$与$-3$，对应的特征向量$\begin{pmatrix} k \\ k \end{pmatrix}$与$\begin{pmatrix} -l \\ l \end{pmatrix}$是正交的．

在例 6 中，$A=\begin{pmatrix} -1 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & -1 \end{pmatrix}$的特征值是 3 与$-6$．$\mathscr{S}_3$的两个生成元 $\mathbf{v}_1$ 与 $\mathbf{v}_2$ 都与 $\mathscr{S}_{-6}$ 的生成元 $\mathbf{v}_3$正交．因此，所有形如 $c\mathbf{v}_1 + d\mathbf{v}_2$ 的向量都与形如 $k\mathbf{v}_3$ 的向量正交．

例 9 验证，埃尔米特矩阵$\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$对应不同特征值的特征向量在 $\mathscr{U}^2$ 中正交．

该矩阵的特征值是 $\lambda=1$ 与 3．特征向量空间 $\mathscr{S}_1$ 与 $\mathscr{S}_3$ 分别由向量

$$\mathbf{v}=\begin{pmatrix} -i \\ 1 \end{pmatrix} \quad 与 \quad \mathbf{w}=\begin{pmatrix} i \\ 1 \end{pmatrix}$$

生成．它们是正交的，因为

$$\langle \mathbf{v}, \mathbf{w} \rangle = \bar{v}_1 w_1 + \bar{v}_2 w_2 = i \cdot i + 1 \cdot 1 = 0.$$

练　　习

5．求下列埃尔米特矩阵的特征值，并且验证，对应不同特征值的特征向量是正交的．

(i) $\begin{pmatrix} 2 & 2+i \\ 2-i & -2 \end{pmatrix}$；(ii) $\begin{pmatrix} 8 & -6 \\ -6 & 17 \end{pmatrix}$；

(iii) $\begin{pmatrix} -1 & 2 & 0 \\ 2 & 0 & -2 \\ 0 & -2 & 1 \end{pmatrix}$；(iv) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix}$；(v) $\begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & 2 \\ -2 & 2 & 0 \end{pmatrix}$．

6．把课文各例题的计算补充完整．

7．已知 S 是实斜称矩阵 ($S'=-S$)，证明，iS 是埃尔米特矩阵．由此推出，S 的非零特征值都是纯虚数，亦即形如 ib 的数，其中 b 是实数．

8．以矩阵

$$S=\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

为例说明练习 7．

9．证明，(非对称)矩阵

$$\begin{pmatrix} 3 & -5 \\ -1 & -1 \end{pmatrix}$$

对应不同特征值的特征向量不是正交的．

10．实斜称矩阵 S 的对应不同特征值的特征向量是否组成正交集？

11．设 A 与 B 是实对称矩阵．证明，$AB-BA$ 是斜称矩阵．设

$$A=\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},\quad B=\begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix},$$

求 $AB-BA$ 的特征值．

15.3　实对称矩阵的对角化

我们现在就要研究关于对称矩阵的主要定理了．我们先证明一个引理，因为在证明主要定理时要用到它．

引理 如果A是$n\times n$对称矩阵，C是任意的$n\times m$矩阵，则$C'AC$是对称矩阵.

证明

$$(C'AC)'=C'A'(C')'=C'AC.$$

后一等号是由于$A'=A$，$(C')'=C$. 由此可见，$C'AC$等于自己的转置.

定理 III 设A是一个实对称矩阵，则存在一个正交矩阵C，使$C^{-1}AC$是对角矩阵.

证明 假定A是$n\times n$矩阵. 如果A有n个不同的特征值，则该定理只是上节定理 II 的简单推论. 所要求的正交矩阵C的列由正规化后的特征向量组成，这些向量是两两正交的. 当A没有n个不同特征值时(此时，A的特征方程有重根)，情况就不那么简单. 下面的证明就是针对这一情况的.

用反证法，假定不存在正交矩阵C，使$C^{-1}AC$是对角矩阵. 此时，不可能找到A的一组特征向量，使它组成$\mathscr{E}^n$(列向量空间)的正规正交基. 不然的话，以这些基向量为列的矩阵C就满足定理的要求.

取尽可能大的正规正交集$\mathbf{c}_1,\cdots,\mathbf{c}_r$，使其中每个向量都是矩阵$A$的特征向量. 这样的集合中的向量个数至少是 1，但总少于n，亦即$1\leqslant r<n$. 用格拉姆-施密特正交化过程，可以把$\mathbf{c}_1,\cdots,\mathbf{c}_r$扩充为$\mathscr{E}^n$的一组正规正交基$\mathbf{c}_1,\cdots,\mathbf{c}_r,\mathbf{d}_{r+1},\cdots,\mathbf{d}_n$. 我们注意到$\mathbf{d}_{r+1},\cdots,\mathbf{d}_n$的任何线性组合都不是$A$的特征向量，因若不然，我们可以找到$r+1$个向量组成的正规正交集，使其中每个向量都是$A$的特征向量，这与$r$的定义矛盾.

再设C是正交矩阵，它的列是

$$\mathbf{c}_1,\cdots,\mathbf{c}_r,\mathbf{d}_{r+1},\cdots,\mathbf{d}_n.$$

我们计算：

$$AC=A(\mathbf{c}_1\cdots\mathbf{c}_r\ \ \mathbf{d}_{r+1}\cdots\mathbf{d}_n)$$

$$=(\lambda_1\mathbf{c}_1\cdots\lambda_r\mathbf{c}_r\ A\mathbf{d}_{r+1}\cdots A\mathbf{d}_n)$$

$$=(\mathbf{c}_1\cdots\mathbf{c}_r\ \mathbf{d}_{r+1}\cdots\mathbf{d}_n)\begin{pmatrix}\lambda_1 & 0 & \cdots & 0 & b_{1\ r+1} & \cdots & b_{1n}\\ 0 & \lambda_2 & \cdots & 0 & b_{2\ r+1} & \cdots & b_{2n}\\ 0 & 0 & \ddots & \vdots & \vdots & & \vdots\\ \vdots & \vdots & & \lambda_r & \vdots & & \vdots\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0 & b_{n\ r+1} & \cdots & b_{nn}\end{pmatrix}$$

$$=CB,$$

这里 $B=(b_{ij})$，其中 $b_{ii}=\lambda_i(i=1,\cdots,r)$，$b_{ij}=0(i\neq j$ 且 $j=1,\cdots,r)$.

因为C是正交矩阵，$C^{-1}=C'$. 所以 $B=C^{-1}AC=C'AC$. 根据引理，B是对称矩阵. 所以 $b_{ji}=0(i\neq j$ 且 $j=1,\cdots,r)$，亦即

$$B=\begin{pmatrix}\lambda_1 & & & & & 0\\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ 0 & & & & & \boxed{B_1}\end{pmatrix},$$

其可能非零的元素在主对角线上以及右下角的 $(n-r)\times(n-r)$ 子矩阵 B_1 中.

B_1本身也是对称矩阵(因为B是对称矩阵)，所以它有一个实特征值μ以及它所对应的实特征向量，设此向量的 $n-r$ 个分量为 $y_{r+1},\cdots,y_n$，即

$$B_1\begin{pmatrix}y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=\mu\begin{pmatrix}y_{r+1}\\ \vdots\\ y_n\end{pmatrix}.$$

因此

$$B\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=\begin{pmatrix}\lambda_1 & & & 0\\ & \ddots & & \\ & & \lambda_r & \\ 0 & & & \boxed{B_1}\end{pmatrix}\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=\mu\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}.$$

因为 $B=C^{-1}AC$,所以

$$AC\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=CB\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=\mu C\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix},\tag{12}$$

其中

$$C\begin{pmatrix}0\\ \vdots\\ 0\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=0\,\mathbf{c}_1+\cdots+0\,\mathbf{c}_n+y_{r+1}\mathbf{d}_{r+1}+\cdots+y_n\mathbf{d}_n(\text{练习}\quad 12)$$
$$=\mathbf{c}_{r+1}.$$

向量 $\mathbf{c}_{r+1}$ 是 $\mathbf{d}_{r+1},\cdots,\mathbf{d}_n$ 的线性组合，并且不是零向量；而根据(12)式，它是 A 的对应特征值 μ 的特征向量．这与早先的断言矛盾．证毕．

定理 III 的推论 已知 A 是 $n\times n$ 实对称矩阵，则存在由 A 的 n 个特征向量组成的正规正交集．

证明 定理 III 所说的矩阵 C 的列就组成这样的集合．

例 10 以 13.2 节例 5 与例 6 的矩阵为实例来说明定理 III．

在例 5 中：特征值是 -1 与 -3. 向量 $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ 与 $\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$ 分别生成 $\mathscr{S}_{-1}$ 与 $\mathscr{S}_{-3}$. 矩阵

$$C=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$$

是正交矩阵，并且

$$C^{-1}\begin{pmatrix}-2 & 1\\ 1 & -2\end{pmatrix}C=\begin{pmatrix}-1 & 0\\ 0 & -3\end{pmatrix}.$$

在例 6 中：特征值是 3 与 -6. $\mathscr{S}_3$ 中的每个向量与 $\mathscr{S}_{-6}$ 中的每个向量都正交. 所以，只要分别找出 $\mathscr{S}_3$ 与 $\mathscr{S}_{-6}$ 的正规正交基，则由这些基向量为列的正交矩阵 C 使原矩阵对角化.

先考虑 $\mathscr{S}_3$：该空间由向量

$$\mathbf{v}_1=\begin{pmatrix}1\\2\\0\end{pmatrix} \quad 与 \quad \mathbf{v}_2=\begin{pmatrix}0\\-2\\1\end{pmatrix}$$

生成. 把 $\mathbf{v}_1$ 正规化，我们得到 $\mathscr{S}_3$ 的正规正交基的第一个向量 $\mathbf{u}_1=\frac{1}{\sqrt{5}}\mathbf{v}_1$. 再取纯量 c，使 $\mathbf{v}_1-c\mathbf{v}_2$ 与 $\mathbf{v}_1$ 正交，亦即使

$$\langle\mathbf{v}_1,\mathbf{v}_1-c\mathbf{v}_2\rangle=5-4c=0.$$

因而，$\mathscr{S}_3$ 的向量 $\mathbf{v}_1-\frac{5}{4}\mathbf{v}_2$ 与 $\mathbf{v}_1$ 正交. 正规化后，我们得到

$$\mathbf{u}_2=\frac{1}{\sqrt{45}}\begin{pmatrix}4\\-2\\5\end{pmatrix}.$$

$\mathbf{u}_1$ 与 $\mathbf{u}_2$ 一起组成了 $\mathscr{S}_3$ 的正规正交基.

$\mathscr{S}_{-6}$ 的正规正交基只由一个向量

$$\mathbf{u}_3=\frac{1}{3}\begin{pmatrix}-2\\1\\2\end{pmatrix}$$

组成．正交矩阵

$$C=\begin{pmatrix}\frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} & -\frac{2}{3}\\ \frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} & \frac{1}{3}\\ 0 & \frac{5}{\sqrt{45}} & \frac{2}{3}\end{pmatrix}\tag{13}$$

使

$$C^{-1}\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}C=\begin{pmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & -6\end{pmatrix}.\tag{14}$$

$\mathscr{S}_3$ 的正规正交基可以有多种取法，我们刚才找到的取法并非最好的．可以看出，$\frac{1}{3}(\mathbf{v}_1+2\mathbf{v}_2)$ 与 $\frac{1}{3}(2\mathbf{v}_1+\mathbf{v}_2)$ 组成 $\mathscr{S}_3$ 的另一组正规正交基，这两个向量与 $\mathbf{u}_3$ 一起组成了正交矩阵

$$C=\frac{1}{3}\begin{pmatrix}1 & 2 & -2\\ -2 & 2 & 1\\ 2 & 1 & 2\end{pmatrix}.\tag{15}$$

这个矩阵在(14)式中所起的作用与矩阵(13)完全一样．

对角化过程的总结

给定一个 $n\times n$ 对称矩阵 A．若要确定一个正交矩阵 C，使 $C^{-1}AC$ 是对角矩阵，则

(i) 求 A 的特征值 $\lambda_1,\cdots,\lambda_r$ 和对应的特征向量空间 $\mathscr{S}_{\lambda_1},\cdots,\mathscr{S}_{\lambda_r}$；

(ii) 对 $i=1,\cdots,r$，把格拉姆-施密特正交化过程用于 $\mathscr{S}_{\lambda_i}$，

作出 $\mathscr{S}_{\lambda_i}$ 的正规正交基；

(iii) 以(ii)中所得到的 $\mathscr{S}_{\lambda_1},\cdots,\mathscr{S}_{\lambda_r}$ 的正规正交基的基向量为列，作出矩阵 C。

假定 $\mathscr{S}_{\lambda_i}$ 的维数为 n_i，则上述理论保证了

$$n_1+\cdots+n_r=n,$$

并且 C 是 $n\times n$ 正交矩阵，使

$$C^{-1}AC=\begin{pmatrix} \lambda_1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & \lambda_r & \\ & & & & & & & \ddots \\ 0 & & & & & & & \lambda_r \end{pmatrix}=D, \tag{16}$$

其中 λ_i 出现 n_i 次。

$C^{-1}AC$ 的特征多项式是

$$|D-\lambda I|=(-1)^n(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_r)^{n_r}. \tag{17}$$

根据 13.3 节的定理 VII，A 与 D 有相同的特征多项式．因此，我们得到如下的结果．

定理 IV 设 λ_i 是实对称矩阵 A 的特征值，则 $\mathscr{S}_{\lambda_i}$ 的维数与因子 $\lambda-\lambda_i$ 在 A 的特征多项式中出现的重数相同．

对于一般的 $n\times n$ 矩阵，该定理不成立(参看 13.2 节的例 7)。

练　　习

12．设 A 为 $n\times n$ 矩阵，它的列是 $\mathbf{a}_1,\cdots,\mathbf{a}_n$。又设 $\mathbf{x}$ 是一个列向量，它的分量是 $x_1,\cdots,x_n$。证明，$A\mathbf{x}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+\cdots+x_n\mathbf{a}_n$。

13．在定理 III 的证明中，我们曾说，向量 $\mathbf{c}_{r+1}$ 是 $\mathbf{d}_{r+1},\cdots,\mathbf{d}_n$ 的线性组合，并且不是零向量．证明这一说法是正确的。

14．对下列各对称矩阵 A，找一个正交矩阵 C，使 $C^{-1}AC$ 是对角矩阵：

(i) $\begin{pmatrix} -1 & 2 & 0 \\ 2 & 0 & -2 \\ 0 & -2 & 1 \end{pmatrix}$；(ii) $\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$；(iii) $\frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$；

(iv) $\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$；(v) $\begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & 2 \\ -2 & 2 & 0 \end{pmatrix}$。

15．证明，除了旋转角是 π 的倍数情况外，旋转矩阵不可能是对称矩阵。

16．证明，一个 3×3 矩阵如果是迹为 1 的对称矩阵，则一定是反射的矩阵(其例子参看练习 14(iii))。

第十六章　二次型的化简

16.1　化简过程

本节中，我们要用十五章的结果来说明，怎样通过适当的坐标变换，把一个二次型化为简单的形式. 其根据是下面这些事实.

定理 I　设Q是 $x_1,\cdots,x_n$ 的二次型，其对称矩阵为A. 又设$\mathbf{x}=C\mathbf{y}$ 定义一个坐标变换. 则Q用新坐标 $y_1,\cdots,y_n$ 表出时，其对称矩阵为 $C'AC$.

证明

$$Q=\mathbf{x}'A\mathbf{x}=(C\mathbf{y})'AC\mathbf{y}=\mathbf{y}'(C'AC)\mathbf{y}.$$

定理 II　设 D 是对角矩阵，其对角元素从上到下依次是 $\lambda_1,\cdots,\lambda_n$，则

$$\mathbf{y}'D\mathbf{y}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2.$$

证明　显然.

定理 III　设A是$n\times n$对称矩阵，其特征多项式为$(-1)^n(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_r)^{n_r}$.又设$Q$是$x_1,\cdots,x_n$的二次型，其矩阵为$A$，则Q可以用坐标变换$\mathbf{x}=C\mathbf{y}$(其中$C$是正交矩阵)化简为二次型

$$\lambda_1y_1^2+\cdots+\lambda_1y_{n_1}^2+\lambda_2y_{n_{1+1}}+\cdots+\lambda_ry_n^2. \tag{1}$$

证明　根据 15.3 节定理 III 后面的讨论，我们可以取正交矩阵C，使

$$C^{-1}AC=D=\begin{pmatrix}\lambda_1 & & & & & & 0\\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \lambda_2 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ 0 & & & & & & \lambda_r\end{pmatrix},$$

其中 λ_i 出现 n_i 次．置 $\mathbf{x}=C\mathbf{y}$，则 Q 用 $y_1,\cdots,y_n$ 表出时，其对称矩阵为 $C'AC=C^{-1}AC=D$，于是，定理的结论从定理 II 得出．

例 1

二次型

$$Q=-x_1^2+2x_2^2-x_3^2+4x_1x_2+8x_1x_3-4x_2x_3 \tag{2}$$

在坐标变换 $\mathbf{x}=C\mathbf{y}$ 下化简为

$$3y_1^2+3y_2^2-6y_3^2, \tag{3}$$

其中 C 是正交矩阵

$$\frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \tag{4}$$

（参看 15.3 节的矩阵(13)）．

简单形式(3)揭示了二次型 Q 的性质．现在，要给出使 Q 取正值、负值和零值的 x_1，x_2，x_3 值的分布已是件容易的事了．事实上

$$Q \gtreqless 0 \quad \text{分别对应于} \quad y_1^2+y_2^2 \gtreqless 2y_3^2.$$

例如

$$\text{当} \quad y_1=y_2=y_3=1 \text{ 时}, \quad Q=0.$$

再从 $\mathbf{x}=C\mathbf{y}$（C 由(4)式定义），我们得到

$$\text{当} \quad x_1=\frac{1}{3}(1+2-2)=\frac{1}{3}, \quad x_2=\frac{1}{3}, \quad x_3=\frac{5}{3} \text{ 时}, \quad Q=0.$$

又如

$$\text{当} \quad y_1=y_2=0, \quad y_3=1 \text{ 时}, \quad Q=-6.$$

这种情况相当于

$$x_1=-\frac{2}{3}, \qquad x_2=\frac{1}{3}, \qquad x_3=\frac{2}{3}.$$

总可以找到 x_1,x_2,x_3 的适当的值，使 Q 取任一指定的数值．

为了看到这一点，我们注意到

$$\text{要 } Q=k>0, \quad \text{只要 } 3y_1^2=k, \quad y_2=y_3=0;$$
$$\text{要 } Q=-k<0, \quad \text{只要 } y_1=y_2=0, \quad 6y_3^2=k.$$

x_1, x_2, x_3 相应的值由关系式 $\mathbf{x}=C\mathbf{y}$ 给出.

定理 III 在许多离不开直角坐标系的几何性质的问题中都是十分重要的.（我们回忆一下，当 C 是正交矩阵时，$\mathbf{x}=C\mathbf{y}$ 定义了从一个直角坐标系到另一个直角坐标系的变换.）参看本节的例 3 与16.2 节.

在与正交性无关的问题中，二次型还能够进一步化简. 在下面的定理中，我们描述了以后需要的一个特殊情况.

定理 IV 设对称矩阵 A 的所有特征值都是正数，Q 是 $x_1, \cdots, x_n$ 的二次型，其矩阵为 A，则存在一个坐标变换 $\mathbf{x}=C\mathbf{z}$，其中 C 是实的非奇异矩阵，但不一定是正交矩阵，可把 Q 化简为

$$z_1^2+\cdots+z_n^2.$$

证明 根据定理 III，存在一个坐标变换 $\mathbf{x}=C_1\mathbf{y}$，把 Q 化简成 (1) 的形式. 再用坐标变换

$$y_1=\frac{1}{\sqrt{\lambda_1}}z_1, \cdots, \quad y_{n_1}=\frac{1}{\sqrt{\lambda_1}}z_{n_1}, \cdots, \quad y_n=\frac{1}{\sqrt{\lambda_r}}z_n,$$

显然把 Q 化简为

$$z_1^2+\cdots+z_n^2.$$

上述坐标变换在实数系范围内能够实现，因为假定了 A 的所有特征值都是正的. 用矩阵记号，我们有 $\mathbf{y}=C_2\mathbf{z}$，其中 C_2 是对角矩阵，它的对角元素依次是：n_1 个 $\frac{1}{\sqrt{\lambda_1}}$，$n_2$ 个 $\frac{1}{\sqrt{\lambda_2}}$，$\cdots$，$n_r$ 个 $\frac{1}{\sqrt{\lambda_r}}$. 置 $C=C_1C_2$，则 $\mathbf{x}=C\mathbf{z}$ 就是所需要的坐标变换.

定理 IV 的推论 设 A, C 如定理 IV，则

$$C'AC=I \text{（恒等矩阵）}.$$

证明 二次型 $z_1^2+\cdots+z_n^2$ 的矩阵是恒等矩阵. 上述结论从定理 I 即得出.

注意,如果 A 有负特征值或零特征值,则实矩阵 C_2 不能作出.

例 2

二次型

$$Q=8x_1^2-12x_1x_2+17x_2^2 \tag{6}$$

通过坐标变换

$$\mathbf{x}=C\mathbf{y}, \quad C=\frac{1}{\sqrt{5}}\begin{pmatrix}2 & -1\\ 1 & 2\end{pmatrix}$$

化简为

$$5y_1^2+20y_2^2. \tag{7}$$

(请验证: Q 的对称矩阵为

$$A=\begin{pmatrix}8 & -6\\ -6 & 17\end{pmatrix};$$

A 的特征值为 5 与 20; 对应的单位特征向量就是 C 的列.)置

$$B=\begin{pmatrix}\frac{1}{\sqrt{5}} & 0\\ 0 & \frac{1}{\sqrt{20}}\end{pmatrix}, \quad C'=CB=\frac{1}{10}\begin{pmatrix}4 & -1\\ 2 & 2\end{pmatrix}.$$

变换 $\mathbf{x}=C'\mathbf{z}$ 把 Q 化简为

$$z_1^2+z_2^2. \tag{8}$$

注意,C' 不是正交矩阵.

从(7)式或(8)式容易看出,当 x_1 与 x_2 取实值时,Q 不能取负值. 事实上,除了 $y_1=y_2=0$,即 $x_1=x_2=0$ 之外,Q 总是正的. 具有这一性质的二次型称为正定的. 进一步的讨论参看 16.3 节.

例 3 主惯性矩

惯性矩阵 J(参看 15.1 节(9) 式)是定义在一个直角坐标系

Ox, Oy, Oz 中的。因为 J 是对称矩阵，所以存在直角坐标系的变换

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = C \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} \qquad (C\text{ 是正交矩阵}),$$

使

$$C^{-1}JC = C'JC = \begin{pmatrix} A^* & 0 & 0 \\ 0 & B^* & 0 \\ 0 & 0 & C^* \end{pmatrix} = J^*. \tag{9}$$

刚体的角速度在新坐标系中由 ω^* 给出，这里 $\omega = C\omega^*$. 在旧坐标系中，刚体的动能是 $T = \frac{1}{2}\omega' J \omega$. 所以

$$\begin{aligned} T &= \frac{1}{2}(C\omega^*)' J C \omega^* \\ &= \frac{1}{2}\omega^{*\prime}(C'JC)\omega^* = \frac{1}{2}\omega^{*\prime} J^* \omega^*. \end{aligned} \tag{10}$$

因而，J^* 是该刚体在新坐标系中的惯性矩阵。在新坐标系中，惯性积都是零。

轴 Ox^*, Oy^*, Oz^* 称为该刚体在 O 点的主轴，而 A^*, B^*, C^* 称为它在 O 点的主惯性矩。注意，主惯性矩只不过是惯性矩阵的特征值。

在主轴组成的坐标系中，动能的表达式最简单：

$$T = \frac{1}{2}(A^*\omega_{x^*}^2 + B^*\omega_{y^*}^2 + C^*\omega_{z^*}^2).$$

例 4 数字实例：主惯性矩

我们来求 15.1 节例 5 中的质点系在 O 点的主轴与主惯性矩。为了简化计算过程，我们仅考虑 $m=1$, $a=b=c=1$(单位正方体)的特殊情况。在所给的坐标系中，

$$J=\begin{pmatrix} 8 & -2 & -2 \\ -2 & 8 & -2 \\ -2 & -2 & 8 \end{pmatrix}.$$

J的特征值是 4 与 10，对应的特征向量空间是

$$\mathscr{S}_4,\ 由\begin{pmatrix}1\\1\\1\end{pmatrix}生成;$$

$$\mathscr{S}_{10},\ 由\begin{pmatrix}1\\0\\-1\end{pmatrix}与\begin{pmatrix}1\\-2\\1\end{pmatrix}生成.$$

据此，我们作出正交矩阵

$$C=\frac{1}{\sqrt{6}}\begin{pmatrix}\sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & -\sqrt{3} & 1\end{pmatrix},$$

使

$$C^{-1}JC=C'JC=\begin{pmatrix}4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10\end{pmatrix}.$$

一组主轴可取在向量

$$\begin{pmatrix}1\\1\\1\end{pmatrix},\ \begin{pmatrix}1\\0\\-1\end{pmatrix},\ \begin{pmatrix}1\\-2\\1\end{pmatrix}$$

（在旧坐标系中表出）的方向上．主惯性矩为 4，10，10．对应主惯性矩 10 的主轴不是唯一确定的．这两条主轴可以取在 $\mathscr{S}_{10}$ 中任意两个正交向量的方向上．

练　　习

1．找一个坐标变换 $\mathbf{x}=C\mathbf{y}$，其中C为正交矩阵，把二次型 $11x_1^2+11x_2^2-$

$18x_1x_2$ 化简为 $\lambda_1y_1^2+\lambda_2y_2^2$的形式。

2. 用坐标变换 $\mathbf{x}=C\mathbf{y}$，其中 C 为正交矩阵，把下列各二次型化简为 $\lambda_1y_1^2+\lambda_2y_2^2+\lambda_3y_3^2$ 的形式。

(i) $-x_1^2+x_3^2+4x_1x_2-4x_2x_3$;

(ii) $3x_1^2+4x_2^2+5x_3^2+4x_1x_2-4x_2x_3$;

(iii) $x_1^2+x_2^2+x_3^2+4x_1x_2-4x_1x_3+4x_2x_3$;

(iv) $x_1^2+x_2^2+x_3^2+x_1x_2-x_1x_3+x_2x_3$;

(v) $x_1x_2-x_1x_3+x_2x_3$。

对每个有可能取值-1 的二次型，求出使该二次型取值-1的 x_1, x_2, x_3 的值。

3. 找一个坐标变换 $\mathbf{x}=C\mathbf{z}$(C 是实矩阵，但不是正交矩阵)，把练习 1 的二次型化简为 $z_1^2+z_2^2$ 的形式。

4. 找一个坐标变换 $\mathbf{x}=C\mathbf{z}$(C 是实矩阵，但不是正交矩阵)，把练习 2 (v) 的二次型化简为 $z_1^2+z_2^2-z_3^2$ 的形式。对练习 2 (i)的二次型进行类似的化简。

5. 在由在 O 点的主轴组成的坐标系中，表出 15.1 节练习 4 所说的单位正立方体的各个角速度。对每一情况，把动能表示成 $T=\frac{1}{2}\omega^{*\prime}J^*\omega^*$ 的形式，并与原先的结果相比较。

6. 一个刚体绕通过 O 点的一条轴旋转，在一个以 O 为原点的直角坐标系中，它的角动量由 $\mathbf{h}=J\omega$ 给出。经过由正交矩阵 C 定义的坐标变换，它的角动量变成 $\mathbf{h}^*$，这里 $\mathbf{h}=C\mathbf{h}^*$。证明$\mathbf{h}^*=J^*\omega^*$。以上一道题中的单位正立方体为例说明之。

16.2 二次方程

例 5 试判定，在直角坐标系中由方程

$$x_1^2-4x_1x_2-2x_2^2=1 \tag{11}$$

所定义的圆锥曲线是什么类型的？

我们把方程(11)写成

$$\mathbf{x}'A\mathbf{x}=1,\quad 其中\quad A=\begin{pmatrix}1 & -2\\ -2 & -2\end{pmatrix}. \tag{12}$$

我们要用关系式 $\mathbf{x}=C\mathbf{y}$ 转到一个新的直角坐标系中来讨论，这里的 C 必须是一个正交矩阵，且使 $C^{-1}AC$ 为对角矩阵. 可以求出，A 的特征值是 $\lambda=-3$ 与 2，并且 $\mathscr{S}_{-3}$ 与 $\mathscr{S}_2$ 分别由向量 $\begin{pmatrix}1\\2\end{pmatrix}$ 与 $\begin{pmatrix}-2\\1\end{pmatrix}$ 生成. 标准化后，我们得到正交矩阵

$$C=\frac{1}{\sqrt{5}}\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix},\quad \text{它使 } C^{-1}AC=\begin{pmatrix}-3 & 0\\ 0 & 2\end{pmatrix}. \tag{13}$$

方程(12)在新坐标系中成为

$$-3y_1^2+2y_2^2=1. \tag{14}$$

这是双曲线，它的主轴与新坐标轴重合. 据 14.1 节的定理 I，主轴方向上的单位向量在旧坐标系中表出则是

$$\frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix} \quad 与 \quad \frac{1}{\sqrt{5}}\begin{pmatrix}-2\\1\end{pmatrix}.$$

该双曲线的草图参看图 86.

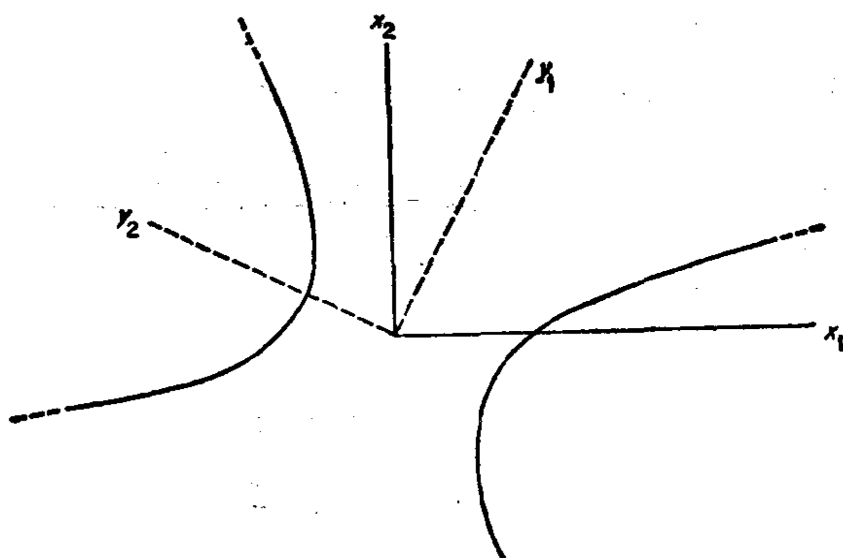


图 86 双曲线 $x_1^2-4x_1x_2-2x_2^2=1$

例 6 讨论圆锥曲线

$$x_1^2-4x_1x_2-2x_2^2+4x_1+4x_2=1. \tag{15}$$

如例 5 那样，用坐标变换 $\mathbf{x}=C\mathbf{y}$，把二次项化简成(14)式左边的平方差．注意

$$x_1=\frac{1}{\sqrt{5}}(y_1-2y_2),\quad x_2=\frac{1}{\sqrt{5}}(2y_1+y_2),$$

我们便得到方程(15)在新坐标系中的表达式

$$-3y_1^2+2y_2^2+\frac{2}{\sqrt{5}}y_1-\frac{4}{\sqrt{5}}y_2=1.$$

因此，配方后得到

$$-3\left(y_1-\frac{2}{\sqrt{5}}\right)^2+2\left(y_2-\frac{1}{\sqrt{5}}\right)^2=1-3\cdot\frac{4}{5}+2\cdot\frac{1}{5}=-1.$$

这是双曲线，中心为 $y_1=\frac{2}{\sqrt{5}}$，$y_2=\frac{1}{\sqrt{5}}$(在新坐标系中表出)．

此例说明了如何用配方法把一次项“吸收”．仅在一种情况下，一次项(比如说 y_1 的项)不能被吸收，那就是 y_1^2 项不出现的时候．此时，所讨论的曲线是抛物线．

圆锥曲线

$$a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2+b_1x_1+b_2x_2+c=0 \tag{16}$$

的类型决定于其二次项的对称矩阵

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{12} & a_{22}\end{pmatrix}$$

的特征值．若不考虑退化的情况，可以把它总结于表 14 中．

再者，A 的特征值是它特征多项式的零点，所以我们有

$$|A-\lambda I|=(\lambda-\lambda_1)(\lambda-\lambda_2).$$

其中 λ_1 与 λ_2 是 A 的特征值(不一定互不相同)．(因为 A 是 2×2 矩阵，λ^2 项的系数为 $+1$.)置 $\lambda=0$，我们便得到

$$|A|=\lambda_1\lambda_2.$$

表 14

A 的特征值	圆锥曲线的类型
同正或同负	椭圆
异号	双曲线
一个特征值为零	抛物线

再与上表比较，我们看到：（非退化的）圆锥曲线(16)是椭圆、双曲线还是抛物线，取决于$|A|$是正数、负数还是零．

例 7 证明，方程

$$9x_1^2+6x_1x_2+x_2^2+2x_1+3x_2+4=0$$

定义的是抛物线．

此时 $A=\begin{pmatrix}9 & 3\\ 3 & 1\end{pmatrix}$，所以$|A|=0$．

练　习

7．判定下列圆锥曲线的类型，并画草图．

(i) $x_1^2-11x_2^2-16x_1x_2=1$；

(ii) $x_1^2-11x_2^2-16x_1x_2+10x_1+10x_2-7=0$；

(iii) $5x_1^2+8x_2^2-4x_1x_2-6x_1-12x_2-36=0$；

(iv) $4x_1^2+x_2^2+4x_1x_2+3x_1+4x_2+1=0$；

(v) $x_1x_2=1$．

8．讨论三维空间的曲面，设它在直角坐标系中的方程是

$$8x^2+8y^2+8z^2-4xy-4xz-4yz=1.$$

（注意：该二次型的对称矩阵就是 16.1 节例 4 的矩阵 J．进行直角坐标系的变换，该曲面的方程可以化简为

$$4x^{*2}+10y^{*2}+10z^{*2}=1.$$

这是椭球面的方程，它称为惯性矩阵为 J 的质点系的矩椭球面．）

16.3 正定二次型

如果一个二次型 $\mathbf{x}'A\mathbf{x}$ 对所有 $\mathbf{x}\neq\mathbf{0}$ 都有 $\mathbf{x}'A\mathbf{x}>0$，则称它是正定的．正定二次型的对称矩阵 A 就称为正定的对称矩阵．

正定二次型经常在物理学问题中出现．例如，一个系统动能的二次型 $\dot{\mathbf{x}}'A\dot{\mathbf{x}}$ 根据其定义的实质就应该是正定的．

定理 V 实对称矩阵 A 是正定的当且仅当它的所有特征值都是正的．

证明 据 16.1 节的定理 III，二次型 $Q=\mathbf{x}'A\mathbf{x}$ 可以用坐标变换 $\mathbf{x}=C\mathbf{y}$ 化简为

$$\lambda_1 y_1^2+\cdots+\lambda_n y_n^2.$$

假定 A 的某个特征值(例如 λ_1)为负值或者零．此时 Q 不是正定的，因为当 $y_1=1, y_i=0(i\neq1)$ 时，Q 的值为 λ_1．反过来，如果对所有的 i，λ_i 都是正的，显然除了 $\mathbf{y}=\mathbf{0}$ 之外，Q 都是正值；而当 $\mathbf{y}=\mathbf{0}$ 时，$\mathbf{x}=C\mathbf{0}=\mathbf{0}$．

例 8 下列对称矩阵中哪些是正定的？

$$\text{(i)}\ \begin{pmatrix}-2 & 1\\ 1 & -2\end{pmatrix};\quad \text{(ii)}\ \begin{pmatrix}9 & 3\\ 3 & 1\end{pmatrix};\quad \text{(iii)}\ \begin{pmatrix}8 & -6\\ -6 & 17\end{pmatrix}.$$

只有(iii)是正定的，它的特征值是 5 和 20．矩阵(i)的特征值为 -1 与 -3，由它所决定的二次型称为负定的，因为此时对所有 $\mathbf{x}\neq\mathbf{0}$，$\mathbf{x}'A\mathbf{x}$ 总是负的．矩阵(ii)的特征值为 0 与 10，由它所决定的二次型称为非负的*，因为此时 $\mathbf{x}'A\mathbf{x}$ 永远不是负数，但是存在非零的 $\mathbf{x}$ 使 $\mathbf{x}'A\mathbf{x}=0$．

15.1 节例 4 中的惯性矩阵是非负的．由于它同动能的联系(参看 15.1 节的(8)式)，其非负性是显然的．事实上，除了质量沿

---

* 非负的(non-negative)二次型又称为半正定的(positive semi-definite)二次型——译者注．

一条直线分布的刚体之外，所有刚体的惯性矩阵都是正定的.

J 的非负性还可以从它的构造法看出．它是从一个形如 $S'S$ 的矩阵经过积分作出的．我们下面证明，凡形如 $S'S$ 的矩阵都是非负的.

定理 VI 设 S 是一个实矩阵，则 $A=S'S$ 是非负的对称矩阵.

证明 显然

$$A'=(S'S)'=S'(S')'=S'S=A,$$

所以 A 是对称矩阵．此外

$$\mathbf{x}'A\mathbf{x}=\mathbf{x}'S'S\mathbf{x}=(S\mathbf{x})'S\mathbf{x}=\mathbf{y}'\mathbf{y},$$

这里 $\mathbf{y}=S\mathbf{x}$．但 $\mathbf{y}'\mathbf{y}=y_1^2+y_2^2+\cdots\geqslant 0$，所以 A 是非负的.

注意，若 $A=S'S$，则

$$|A|=|S'||S|=|S|^2.$$

由此可见，A 是非奇异的当且仅当 S 是非奇异的.

定理 VI 的逆定理也成立(练习 11).

练　　习

9．下列二次型中哪些是正定的？哪些是负定的？哪些是非负的？

(i) $11x_1^2+11x_2^2-18x_1x_2$;

(ii) $4x_1^2+x_2^2-4x_1x_2$;

(iii) x_1x_2;

(iv) $-2x_1^2-5x_2^2+4x_1x_2$.

10．16.1 节练习 2 的二次型中哪些是正定的？哪些是负定的？哪些是非负的？

11．设 A 是非负的对称矩阵.

(i) 证明，存在一个正交矩阵 C，使 $C^{-1}AC=D^2$(一个实对角矩阵的平方).

(ii) 由此推出 $A=S^2$，这里 S 是实对称矩阵．这说明了定理 VI 的逆定理也成立，因为 S 是对称的，$S^2=S'S$.

12. 以下列矩阵为例，说明练习 11：

$$\text{(i)}\begin{pmatrix}29 & -14\\ -14 & 8\end{pmatrix};\quad \text{(ii)}\begin{pmatrix}9 & 3\\ 3 & 1\end{pmatrix}.$$

13. 证明，对称矩阵

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

是正定的当且仅当

$$a_{11}>0 \quad 且 \quad \begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0.$$

16.4 两个二次型的同时化简

在经典动力学中遇到的许多问题常常是由两个二次型描述的，并且其中至少有一个二次型正定．这两个二次型就是一个动力学系统的动能与势能，它们分别可用广义坐标 $x_1, \cdots, x_n$ 表为 $\dot{\mathbf{x}}K\dot{\mathbf{x}}$（正定）与 $\mathbf{x}'P\mathbf{x}$．实例可以参看 15.1 节的例 1.

下面的定理说明了，可以适当选取坐标系，把这样两个二次型同时化为很简单的表达式．在把拉格朗日运动方程应用于一个系统并且分析该系统的特性之前，这种化简是重要的预备步骤．

定理 VII 设 $\mathbf{x}'K\mathbf{x}$ 与 $\mathbf{x}'A\mathbf{x}$ 是 $x_1, \cdots, x_n$ 的二次型，K 正定，则存在一个坐标变换 $\mathbf{x}=C\mathbf{y}$，把这两个二次型同时分别化简为

$$y_1^2+\cdots+y_n^2$$

与

$$\mu_1 y_1^2+\cdots+\mu_n y_n^2,$$

其中 $\mu_1, \cdots, \mu_n$ 是 $|A-\mu K|$ 的零点（不一定互不相同）.

证明 该证明分为两步.

第一步 根据 16.1 节定理 IV 的推论，存在一个矩阵 C_1，使 $C_1'KC_1=I$．置 $C_1'AC_1=A_1$，则 A_1 还是对称矩阵.

第二步 根据 15.3 节定理 III，存在一个正交矩阵 C_2，使

$$C_2^{-1}A_1C_2=D=\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix}.$$

现在令 $C=C_1C_2$. 记住 $C_2^{-1}=C_2'$,容易验证

$$\begin{aligned}C'KC&=I,\\ C'AC&=D.\end{aligned}\tag{17}$$

所以,坐标变换 $\mathbf{x}=C\mathbf{y}$ 把这两个二次型化简为所要求的形式.

最后,从关系式(17)得到

$$C'(A-\mu K)C=\begin{pmatrix}\mu_1-\mu & & 0\\ & \ddots & \\ 0 & & \mu_n-\mu\end{pmatrix}.\tag{18}$$

由此可见,$|A-\mu K|=0$ 当且仅当 $\mu=\mu_i$(对某个 i). 证毕.

我们分两步证明了能起同时化简作用的矩阵 C 的存在性. 事实上,基于(18)式,C 可以用更直接的方法作出来,下面就来证明这一点.

定理 VIII 定理 VII 提到的矩阵 C 的第 i 列 $\mathbf{c}_i$ 是满足

$$A\mathbf{c}_i=\mu_iK\mathbf{c}_i\quad(i=1,\cdots,n)\tag{19}$$

的非零向量.

证明 在(18)式中令 $\mu=\mu_i$,并把 C 写成 $(\mathbf{c}_1\cdots\mathbf{c}_i\cdots\mathbf{c}_n)$. 右边对角矩阵的第 i 列是 $\mathbf{0}$,而左边的第 i 列是 $C'(A-\mu_iK)\mathbf{c}_i$. 因而

$$C'(A-\mu_iK)\mathbf{c}_i=\mathbf{0}.$$

我们知道 C' 是非奇异的,可以用它的逆矩阵乘上式两边,我们就得到了方程(19).

用上面的定理可以把 C 的列确定出来,至多差一个纯量因子. 对于大多数实际应用问题,这已经够了. 一旦需要,纯量因子也不难求出. 我们现在举例说明之.

例 9 找一个坐标变换 $\mathbf{x}=C\mathbf{y}$,把二次型

$$\mathbf{x}'K\mathbf{x}=18x_1^2+4x_1x_2+2x_2^2$$

与

$$\mathbf{x}'A\mathbf{x}=5x_1^2+2x_1x_2+x_2^2$$

同时分别化简为 $y_1^2+y_2^2$ 与 $\mu_1y_1^2+\mu_2y_2^2$.

我们有

$$K=\begin{pmatrix}18 & 2\\ 2 & 2\end{pmatrix},\quad A=\begin{pmatrix}5 & 1\\ 1 & 1\end{pmatrix}.$$

容易验证K是正定的. 再者

$$|A-\mu K|=\begin{vmatrix}5-18\mu & 1-2\mu\\ 1-2\mu & 1-2\mu\end{vmatrix}=(1-2\mu)(4-16\mu).$$

所以, $|A-\mu K|$的零点是 $\mu_1=\frac{1}{2}$, $\mu_2=\frac{1}{4}$.

现在求方程组 $A\mathbf{z}=\mu_iK\mathbf{z}$ 的非平凡解 $(i=1,2)$. 当 $i=1$ 时

$$5z_1+z_2=\frac{1}{2}(18z_1+2z_2)=9z_1+z_2,$$

$$z_1+z_2=\frac{1}{2}(2z_1+2z_2),$$

其通解为 $z_1=0$, $z_2=a$(任意).

当 $i=2$ 时

$$5z_1+z_2=\frac{1}{4}(18z_1+2z_2),$$

$$z_1+z_2=\frac{1}{4}(2z_1+2z_2),$$

其通解为 $z_1=b$, $z_2=-b$(b 任意).

所以,根据我们的定理,存在一组常数 a,b,使坐标变换

$$\mathbf{x}=C\mathbf{y},\quad C=\begin{pmatrix}0 & b\\ a & -b\end{pmatrix} \tag{20}$$

把 $\mathbf{x}'K\mathbf{x}$ 与 $\mathbf{x}'A\mathbf{x}$ 同时分别化简为 $y_1^2+y_2^2$ 与 $\frac{1}{2}y_1^2+\frac{1}{4}y_2^2$. 常数 a

与 b 必须满足 $C'KC=I$. 由此得出 $a=\frac{1}{\sqrt{2}}$, $b=\frac{1}{4}$.

例 10 例 9 在动力学中的应用

设一个动力学系统的动能与势能*分别由二次型

$$T=\dot{\mathbf{x}}'K\dot{\mathbf{x}}=18\,\dot{x}_1^2+4\,\dot{x}_1\dot{x}_2+2\,\dot{x}_2^2$$

与

$$V=\mathbf{x}'A\mathbf{x}=5\,x_1^2+2\,x_1x_2+x_2^2$$

给出. 描述它的运动.

根据例 9,存在一个坐标变换 $\mathbf{x}=C\mathbf{y}$,使

$$T=\dot{y}_1^2+\dot{y}_2^2,$$
$$V=\mu_1y_1^2+\mu_2y_2^2,$$

其中 $\mu_1=\frac{1}{2}$, $\mu_2=\frac{1}{4}$. 用拉格朗日运动方程

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}_i}\right)-\frac{\partial L}{\partial y_i}=0 \quad (i=1,2),$$

其中 $L=T-V$ 为该系统的拉格朗日算子. 由此立刻得出方程组

$$\begin{aligned}\ddot{y}_1+\mu_1y_1=0,\\ \ddot{y}_2+\mu_2y_2=0.\end{aligned} \tag{21}$$

方程组(21)的通解是

$$y_1=c\sin(\sqrt{\mu_1}\,t+\alpha), \quad y_2=d\sin(\sqrt{\mu_2}\,t+\beta).$$

再从(20)式得到

$$x_1=by_2, \quad x_2=ay_1-by_2,$$

其中 $a=\frac{1}{\sqrt{2}}$, $b=\frac{1}{4}$. 由此便得出了在旧坐标系中表示出来的通解. 两类正规振动方式分别对应 $y_2=0, y_1\neq0$ 与 $y_1=0, y_2\neq0$ 这两种情况;对旧坐标系而言,也就是

---

* 参看 Rutherford, *Classical Mechanics*, Third Edition, Oliver & Boyd, 1964, § 92.

$$x_1=0,\ x_2=ac\sin(\sqrt{\mu_1}\,t+\alpha)$$

与

$$x_1=bd\sin(\sqrt{\mu_2}\,t+\beta)=-x_2$$

这两种情况．请注意 a、b 的值在解中是无关紧要的．

练　习

14．对下列每组二次型，找一个坐标变换 $\mathbf{x}=C\mathbf{y}$，把其中两个二次型同时分别化简为 $y_1^2+y_2^2$ 与 $\mu_1y_1^2+\mu_2y_2^2$．每组中先写的二次型是正定的．

(i) $x_1^2-2x_1x_2+4x_2^2$ 与 $-4x_1x_2$；

(ii) $2x_1^2-10x_1x_2+13x_2^2$ 与 $5x_1^2-22x_1x_2+25x_2^2$．

15．(卢瑟福)一个振动系统的动能与势能分别是

$T=ma^2(2\dot{x}_1^2-10\dot{x}_1\dot{x}_2+13\dot{x}_2^2)$，$V=mag(5x_1^2-22x_1x_2+25x_2^2)$．

证明，它的两类正规振动方式的周期分别为 $2\pi\sqrt{a/g}$ 与 $\pi\sqrt{a/g}$．求对应的正规坐标．

16．解 15.1 节练习 1 与练习 2 的振动问题．

第十七章　特征值的计算

17.1　用迭代法求主特征值

矩阵的特征值可以用解特征方程的方法求得．这就要先进行行列式的计算(如果阶数较大，这一步就可能十分冗长)，接着还要求出多项式的零点．我们现在要介绍另外一个便于计算的方法，用此方法能够求出绝对值最大的特征值和对应的特征向量．这个方法适用于绝大多数对称矩阵，也能用于一大类非对称矩阵．

设 A 是 $n\times n$ 矩阵,其特征值为 $\lambda_1,\cdots,\lambda_n$,且

$$|\lambda_1|>|\lambda_2|\geqslant|\lambda_3|\geqslant\cdots\geqslant|\lambda_n|. \tag{1}$$

再假定 A 有 n 个线性无关的特征向量 $\mathbf{v}_1,\cdots,\mathbf{v}_n$,分别对应特征值 $\lambda_1,\cdots,\lambda_n$．因为 $\mathbf{v}_1,\cdots,\mathbf{v}_n$ 生成了 n 维列向量空间，所以其中任意一个向量 $\mathbf{x}$ 可以表为如下的线性组合

$$\mathbf{x}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n.$$

因而

$$A\mathbf{x}=c_1A\mathbf{v}_1+\cdots+c_nA\mathbf{v}_n=c_1\lambda_1\mathbf{v}_1+\cdots+c_n\lambda_n\mathbf{v}_n.$$

更一般地说

$$\begin{aligned}A^r\mathbf{x}&=c_1\lambda_1^r\mathbf{v}_1+\cdots+c_n\lambda_n^r\mathbf{v}_n\\&=\lambda_1^r\left(c_1\mathbf{v}_1+\sum_{i=2}^n\left(\frac{\lambda_i}{\lambda_1}\right)^r c_i\mathbf{v}_i\right).\end{aligned} \tag{2}$$

再由(1)式

$$\text{当}\quad r\to\infty\ \text{时}\quad \left(\frac{\lambda_i}{\lambda_1}\right)\to 0\quad (i=2,\cdots,n).$$

因此，我们看到，只要 $c_1 \neq 0$，

$$\left(\frac{1}{\lambda_1^r c_1}\right) A^r \mathbf{x}$$

收敛于特征向量 $\mathbf{v}_1$.

我们称 λ_1 为 A 的主特征值，对应的特征向量 $\mathbf{v}_1$ 称为主特征向量*.

根据上面的观察，我们可以得到一种迭代法，用以求 A 的主特征值：依次计算

$$A\mathbf{x}, \quad A(A\mathbf{x}) = A^2\mathbf{x}, \quad A(A^2\mathbf{x}) = A^3\mathbf{x}, \quad \cdots.$$

只要 $c_1 \neq 0$，此序列中的向量就越来越接近于 $\mathbf{v}_1$ 的纯量倍. 这些向量的分量有可能变得太大或太小，因此，这一迭代法通常与一个"正规化"过程**结合在一起. 所谓"正规化"过程，就是把一个向量 $\mathbf{x}$ 变为它的纯量倍 $\mathbf{y}$，使 $\mathbf{y}$ 的各分量中绝对值最大者为 ± 1. 这是通过关系式

$$\mathbf{y} = \frac{1}{\max(\mathbf{x})}\mathbf{x} \tag{3}$$

得到的，其中 $\max(\mathbf{x})$ 表示 $\mathbf{x}$ 的各分量中绝对值最大者. 例如，若已知

$$\mathbf{x} = \begin{pmatrix} -2 \\ 2 \\ -4 \end{pmatrix},$$

则

$$\max(\mathbf{x}) = 4, \quad \mathbf{y} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -1 \end{pmatrix}.$$

---

* 主特征值(dominant eigenvalue)与主特征向量(dominant eigenvector)又可译作优势特征值与优势特征向量——译者注.

** 该过程不要与把向量长度变为 1 的正规化过程混淆.

17.2 迭代过程

取任意向量 $\mathbf{x}$，求 $\mathbf{x}_1=A\mathbf{x}$；像上文所说的那样，正规化 $\mathbf{x}_1$，得到 $\mathbf{y}_1$；再求 $\mathbf{x}_2=A\mathbf{y}_1$，并正规化，得到 $\mathbf{y}_2$. 如此续行，作出两个向量序列 $\mathbf{y}_1,\mathbf{y}_2,\cdots,\mathbf{y}_i,\cdots$ 与 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_i,\cdots$，它们之间的关系是

$$\mathbf{x}_i=A\mathbf{y}_{i-1},\qquad \mathbf{y}_i=\frac{1}{\max(\mathbf{x}_i)}\mathbf{x}_i. \tag{4}$$

只要在(2)式中 $c_1\neq 0$，对充分大的 i，我们就有

$$\mathbf{x}_i=\lambda_1\mathbf{y}_{i-1}\quad（近似）\tag{5}$$

与

$$\mathbf{y}_i=\frac{1}{|\lambda_1|}\mathbf{x}_i. \tag{6}$$

因此

$$当\ i\to\infty\ 时,\qquad \max(\mathbf{x}_i)\to|\lambda_1|, \tag{7}$$

并且当 i 充分大时，$\mathbf{y}_i$ 近似于主特征向量. 再从(5)式与(6)式得

$$|\lambda_1|\mathbf{y}_i=\lambda_1\mathbf{y}_{i-1}\quad（近似），$$

所以，λ_1 是正的还是负的，取决于对充分大的 i，$\mathbf{y}_i$ 与 $\mathbf{y}_{i-1}$ 是近似相等还是相差一个负号.

例 1 求矩阵

$$A=\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}$$

的主特征值与主特征向量.

从向量 $\mathbf{x}=\begin{pmatrix}1\\0\\0\end{pmatrix}$ 开始，作出(4)的两个序列. 取四位有效数字，列成下表：

表 15

i	1	2	3	4	5	6
$\mathbf{x}_i$	$\begin{pmatrix}-1\\2\\4\end{pmatrix}$	$\begin{pmatrix}5.250\\-1.500\\-3.000\end{pmatrix}$	$\begin{pmatrix}-3.857\\2.571\\5.143\end{pmatrix}$	$\begin{pmatrix}5.750\\-2.500\\-5.000\end{pmatrix}$	$\begin{pmatrix}-5.348\\2.870\\5.739\end{pmatrix}$	$\begin{pmatrix}5.932\\-2.864\\-5.727\end{pmatrix}$
$\mathbf{y}_i$	$\begin{pmatrix}-0.25\\0.5\\1\end{pmatrix}$	$\begin{pmatrix}1\\-0.2857\\-0.5714\end{pmatrix}$	$\begin{pmatrix}-0.7500\\0.5000\\1\end{pmatrix}$	$\begin{pmatrix}1\\-0.4348\\-0.8696\end{pmatrix}$	$\begin{pmatrix}-0.9318\\0.5000\\1\end{pmatrix}$	$\begin{pmatrix}1\\-0.4828\\-0.9655\end{pmatrix}$

若进一步作迭代，可以看出 $\max(\mathbf{x}_i)\to 6$，且当 $k\to\infty$ 时，

$$\mathbf{y}_{2k}\to\mathbf{v}_1=\begin{pmatrix}1\\-\frac{1}{2}\\-1\end{pmatrix},$$

$$\mathbf{y}_{2k+1}\to-\mathbf{v}_1.$$

A 的主特征值是 -6，$\mathbf{v}_1$ 为其主特征向量。这与 13.2 节例 6 中精确计算的结果是一致的。

加快收敛

从 (1) 式与 (2) 式我们看到，迭代过程的收敛速度依赖于比值 $\left|\frac{\lambda_2}{\lambda_1}\right|$。如果 $\left|\frac{\lambda_2}{\lambda_1}\right|$ 很接近 1，则收敛得很慢。在例 1 中，$\lambda_2=3$，$\left|\frac{\lambda_2}{\lambda_1}\right|=\frac{1}{2}$。即使这样小的比值，在六步迭代以后，对主特征值，我们还只能得到一位有效数字，而无法得到精确度更高的值！有一个方法可以获得更快的收敛速度，它基于如下的简单事实：如果 $\mathbf{v}$ 是 A 的特征向量，对应特征值 λ，则

$$(A-pI)\mathbf{v}=(\lambda-p)\mathbf{v}, \tag{9}$$

这说明了 $\mathbf{v}$ 也是 $A-pI$ 的特征向量．而对应的特征值为 $\lambda-p$．据此，如果在例 1 中我们用特征值是 -10 与 -1 的矩阵 $A-4I$ 作

迭代，就会得出 $A-4I$ 的主特征向量（对应特征值 -10），它同时也是 A 的主特征向量（对应特征值 -6）. 现在，收敛过程受比值 $\frac{1}{10}$ 支配，所以将比原先的过程要快. 下面的计算说明了这一点.

例 2

用矩阵

$$A-4I=\begin{pmatrix}-5 & 2 & 4\\ 2 & -2 & -2\\ 4 & -2 & -5\end{pmatrix}$$

作迭代，也从向量

$$\mathbf{x}=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

开始. 列表如下：

表 10

i	1	2	3	4
$\mathbf{x}_i$	$\begin{pmatrix}-5\\2\\4\end{pmatrix}$	$\begin{pmatrix}-9\\4.4\\8.8\end{pmatrix}$	$\begin{pmatrix}9.889\\-4.933\\-9.867\end{pmatrix}$	$\begin{pmatrix}-9.989\\4.993\\9.987\end{pmatrix}$
$\mathbf{y}_i$	$\begin{pmatrix}-1\\0.4\\0.8\end{pmatrix}$	$\begin{pmatrix}-1\\0.4889\\0.9778\end{pmatrix}$	$\begin{pmatrix}1\\-0.4989\\-0.9978\end{pmatrix}$	$\begin{pmatrix}-1\\0.4999\\0.9998\end{pmatrix}$

在这里，迭代过程的收敛（到主特征值 -10 与主特征向量 $\mathbf{v}_1$）就比例 1 要快得多了.

实际应用中，如果关于矩阵 A 的特征值无任何信息可供利用，则可从用 A 作迭代的过程中得到对 λ_1 与 λ_2 的估计. 如果收敛得太慢，就选取 p，使得用 $A-pI$ 作迭代时，速度能加快一些.

迭代过程的失效

(i) 该过程依赖于选择 $\mathbf{x}$ 使公式(2)中的 $c_1 \neq 0$. 假如从向量

$\mathbf{x}=\begin{pmatrix}1\\0\\1\end{pmatrix}$开始,用例 1 的矩阵 A 作迭代,则得到 $\mathbf{y}_i=\mathbf{x}$(对所有 i),

并且 $\lambda=3$. 原因是 $\mathbf{x}$ 为 A 对应于特征值 3 的征特向量. 为了保证迭代过程得出主特征值,必须用若干个初始向量 $\mathbf{x}$ 去试试.

(ii) 如果违背了假设(1),迭代过程可能失效. 例如,考虑

$$A=\begin{pmatrix}3 & 4\\4 & -3\end{pmatrix},\quad \text{初始向量 } \mathbf{x}=\begin{pmatrix}1\\0\end{pmatrix}. \tag{10}$$

用迭代法得出

$$\mathbf{y}_1=\begin{pmatrix}0.75\\1\end{pmatrix},\qquad \mathbf{y}_2=\begin{pmatrix}1\\0\end{pmatrix},$$

$$\mathbf{y}_3=\mathbf{y}_5=\mathbf{y}_7=\cdots=\mathbf{y}_1,\quad \mathbf{y}_4=\mathbf{y}_6=\mathbf{y}_8=\cdots=\mathbf{y}_2.$$

所以,该过程不会收敛于 A 的特征向量. 在这里,失效的原因是 A 的特征值为 $\lambda_1=5,\lambda_2=-5$,从而 $|\lambda_1|=|\lambda_2|$,违背了假设(1).

补救的方法很简单. 把用 A 迭代换成用 $A-pI$ 迭代,就能够把两个特征值分开了. 取 $p=4.9$,我们就得到一个矩阵,它的特征值是 $5-4.9=0.1$ 与 $-5-4.9=-9.9$. 用 $A-4.9I$ 作迭代,则收敛速度非常快,因为其过程受比值$\frac{1}{99}$支配.

对于特征值满足 $\lambda_1=\lambda_2=\cdots=\lambda_r$, $|\lambda_1|>|\lambda_i|$(对所有 $i>r$)的 $n\times n$ 矩阵,不会产生刚才所说的问题. 公式(2) 可以稍加修改,就把这种情况包括进去了.

17.3 其余特征值的计算

我们仅限于观察某些事实并列出一些参考书.

(i) 假定 A 的主特征值 λ_1 是正数,而其余特征值按其大小顺

序(符号考虑在内)排列为

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n,$$

则最小的特征值 λ_n 可由考虑矩阵 $A-pI$ 而求得，只要把 p 取得使 $\lambda_n - p$ 为 $A-pI$ 的主特征值．这总是可能的．例如，如果 A 的特征值为 5,3,−1 与 −2，则 $A-2I$ 的特征值为 3,1,−3 与 −4，其中 −4 是主特征值．

如果 A 的主特征值是负的，则同样的方法可用于 $-A$．

(ii) A 的中间几个特征值可以用"压缩"法求得．这一技巧就是把 A 换成另一个矩阵，使它的特征值与 A 的特征值除一个外都相同，这个例外是 A 的主特征值换成了零．下面的定理就对称矩阵说明这一方法．

定理 I 设 A 是 $n\times n$ 对称矩阵，它的特征值为 $\lambda_1,\cdots,\lambda_n$．设 $\mathbf{u}_1$ 是对应 λ_1 的单位特征向量，因而 $\mathbf{u}_1^t\mathbf{u}_1=1$，$A\mathbf{u}_1=\lambda_1\mathbf{u}_1$．此时，矩阵

$$B=A-\lambda_1\mathbf{u}_1\mathbf{u}_1^t$$

的特征值是 $0,\lambda_2,\cdots,\lambda_n$．

注意：$\mathbf{u}_1\mathbf{u}_1^t$ 是一个列向量($n\times 1$ 矩阵）与一个行向量（$1\times n$ 矩阵)的乘积，根据矩阵乘法的规则，它是 $n\times n$ 矩阵．

定理 I 的证明 由 15.3 节得知，可以找到由 A 的 n 个特征向量组成的正规正交集 $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_n$，对应的特征值分别为 $\lambda_1,\lambda_2,\cdots,\lambda_n$(这些特征值不一定互不相同)．此时，

$$\begin{aligned}B\mathbf{u}_i &= A\mathbf{u}_i-\lambda_1\mathbf{u}_1\mathbf{u}_1^t\mathbf{u}_i\\ &=\lambda_i\mathbf{u}_i-\lambda_1\mathbf{u}_1(\mathbf{u}_1^t\mathbf{u}_i).\end{aligned}$$

但 $i=1$ 时，$\mathbf{u}_1^t\mathbf{u}_i=1$；在其他情况下 $\mathbf{u}_1^t\mathbf{u}_i=0$（据正交性)．所以

$$B\mathbf{u}_i=\begin{cases}0, & \text{当 } i=1 \text{ 时；}\\ \lambda_i\mathbf{u}_i, & \text{当 } i\neq 1 \text{ 时．}\end{cases}$$

这样，就证明了 B 的特征值是 $0,\lambda_2,\cdots,\lambda_n$．注意 A 的特征向量

集合 $\mathbf{u}_1,\cdots,\mathbf{u}_n$ 同时也是 B 的特征向量集合．对于 A，$\mathbf{u}_1$ 对应的特征值是 λ_1；而对于 B，$\mathbf{u}_1$ 对应的特征值是零．

例 3

矩阵

$$A=\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}$$

的单位主特征向量是

$$\mathbf{u}_1=\frac{1}{3}\begin{pmatrix}2\\ -1\\ -2\end{pmatrix},$$

它对应主特征值 $\lambda_1=-6$．我们作

$$\mathbf{u}_1\mathbf{u}_1^t=\frac{1}{9}\begin{pmatrix}2\\ -1\\ -2\end{pmatrix}(2\quad -1\quad -2)=\frac{1}{9}\begin{pmatrix}4 & -2 & -4\\ -2 & 1 & 2\\ -4 & 2 & 4\end{pmatrix}.$$

所以

$$B=\begin{pmatrix}-1 & 2 & 4\\ 2 & 2 & -2\\ 4 & -2 & -1\end{pmatrix}-(-6)\frac{1}{9}\begin{pmatrix}4 & -2 & -4\\ -2 & 1 & 2\\ -4 & 2 & 4\end{pmatrix}$$

$$=\frac{1}{3}\begin{pmatrix}5 & 2 & 4\\ 2 & 8 & -2\\ 4 & -2 & 5\end{pmatrix}.$$

B 的特征值是 0 与 3（两重）．

如欲得到有关特征值计算的更深入的论述，我们建议读者参阅如下书籍：

1．B. Noble, *Numerical Methods*, 1, *Iteration, Programming and Algebraic Equations*, Oliver & Boyd, 1964.

2．J. H. Wilkinson, *The Algebraic Eigenvalue Problem*,

Clarendon Press, Oxford, 1965.

练　　习

1．用迭代法求矩阵

$$A=\begin{pmatrix} 11 & -9 \\ -9 & 11 \end{pmatrix}$$

的主特征值．用压缩法求其第二个特征值．

2．迭代法对矩阵

$$A=\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$$

失效．把迭代法用于矩阵 $A-4.9\,I$，求出 A 的一个特征值及其对应的特征向量．

3．用迭代法和压缩法求矩阵

$$\begin{pmatrix} 4 & 2 & 4 \\ 2 & 7 & -2 \\ 4 & -2 & 4 \end{pmatrix}$$

的特征值．

第十八章　代数学的若干应用

18.1　分子振动：CO_2 分子

求一个振动分子的固有频率的问题实质上是一个特征值问题．作为这一课题的一个引论，我们在这里考虑二氧化碳的分子．这是一个线性分子，亦即，当这一分子处于平衡状态时，它的所有原子都在一条直线上．

我们先考虑纵向振动．其计算类似于在 13.1 节例 3 中提出而在 13.3 节解决的弹簧的振动问题．我们假定这一分子的所有运动都是很小的．

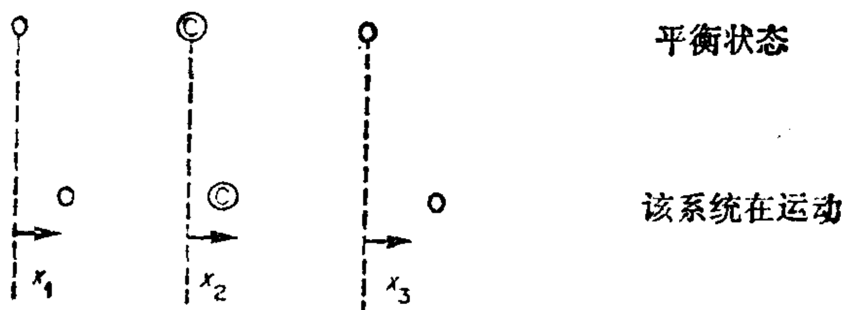


图 87　CO_2 分子的纵向运动

设 x_1 与 x_3 表示当这一系统作纵向振动时，两个氧原子偏离平衡位置的位移，而 x_2 表示碳原子偏离平衡位置的位移．假定恢复力是偏离平衡位置的位移的线性函数，并假定碳原子和一个氧原子之间的力常数（相当于弹簧的弹性系数）为 k，而两个氧原子之间的力常数为 k'．一个氧原子的质量用 m 表示，碳原子的质量用 M 表示，则运动方程是

$$\left.\begin{aligned}m\ddot{x}_1&=k(x_2-x_1)+k'(x_3-x_1),\\M\ddot{x}_2&=-k(x_2-x_1)+k(x_3-x_2),\\m\ddot{x}_3&=-k(x_3-x_2)-k'(x_3-x_1).\end{aligned}\right\}\tag{1}$$

这个方程组可以用矩阵记号写成

$$\ddot{\mathbf{x}}=A\mathbf{x},\text{这里 } A=\begin{pmatrix}-(p+q) & p & q\\ r & -2r & r\\ q & p & -(p+q)\end{pmatrix},\tag{2}$$

其中

$$p=\frac{k}{m},\quad q=\frac{k'}{m},\quad r=\frac{k}{M}.$$

为了解这个方程组,我们着手找一个矩阵 C,使 $C^{-1}AC$ 是对角矩阵.然后,用坐标变换 $\mathbf{x}=C\mathbf{y}$ 就会把正规方式揭示出来.

A 的特征方程是

$$|A-\lambda I|=-\lambda(\lambda+p+2q)(\lambda+p+2r)=0.\tag{3}$$

所以 A 的特征值是 $\lambda=0,-(p+2q),-(p+2r)$.对应的特征向量空间的生成元如下所示:

$$\mathscr{S}_0:\begin{pmatrix}1\\1\\1\end{pmatrix};\quad \mathscr{S}_{-(p+2q)}:\begin{pmatrix}1\\0\\-1\end{pmatrix};\quad \mathscr{S}_{-(p+2r)}:\begin{pmatrix}1\\-\dfrac{2m}{M}\\1\end{pmatrix}.$$

再置

$$\mathbf{x}=C\mathbf{y},\quad \text{这里 } C=\begin{pmatrix}1 & 1 & 1\\ 1 & 0 & -\dfrac{2m}{M}\\ 1 & -1 & 1\end{pmatrix},\tag{4}$$

则

$$C^{-1}AC=\begin{pmatrix}0 & 0 & 0\\ 0 & -(p+2q) & 0\\ 0 & 0 & -(p+2r)\end{pmatrix}.\tag{5}$$

运动方程化简为 $\ddot{y}=(C^{-1}AC)y$，亦即

$$\left.\begin{aligned}\ddot{y}_1&=0,\\ \ddot{y}_2+(p+2q)y_2&=0,\\ \ddot{y}_3+(p+2r)y_3&=0.\end{aligned}\right\}\tag{6}$$

其通解为

$$\left.\begin{aligned}y_1&=c_1t+d,\\ y_2&=c_2\sin(\sqrt{p+2q}\ t+\alpha),\\ y_3&=c_3\sin(\sqrt{p+2r}\ t+\beta).\end{aligned}\right\}\tag{7}$$

三类正规方式分别是当 $i=1,2,3$ 时由 $y_i\neq0, y_j=0(j\neq i)$ 所给出的．由(4)式，这三类正规方式分别对应于

$$\left.\begin{aligned}&\text{(i)}\quad x_1=x_2=x_3=c_1t+d,\\ &\text{(ii)}\quad x_2=0, x_1=-x_3=c_2\sin(\sqrt{p+2q}\,t+\alpha),\\ &\text{(iii)}\quad x_1=x_3=c_3\sin(\sqrt{p+2r}\,t+\beta),\quad x_2=-\left(\frac{2m}{M}\right)x_1\ .\end{aligned}\right\}\tag{8}$$

第一类正规方式（对应特征值 $\lambda=0$）只是一个平移．其余两类正规方式都是在平衡位置附近的振动（图 88）．其周期分别是

$$\frac{2\pi}{\sqrt{p+2q}} \text{ 与 } \frac{2\pi}{\sqrt{p+2r}}.$$



图 88　CO_2 分子纵向运动的正规方式

我们在解这个纵向振动问题时，先从方程组(1)开始．这个方程组是根据如下定律列出的：力＝质量×加速度．在更复杂的问题中还要用到拉格朗日运动方程．我们曾在 16.4 节例 10 中用过这个方法．

不管用什么方法，要确定振动频率，总要牵涉到把一个行列式

因式分解. 显然,对于高阶行列式,这可能是一个十分冗长的过程. (3) 式中$|A-\lambda I|$的因式分解,尽管仅仅是三阶行列式,其计算过程已是繁得令人讨厌了（文中我们已把这些计算略去）. 而若在三维空间讨论 CO_2 分子的运动，其方程组要涉及九个坐标，每个原子三个坐标，问题就归结到求一个九阶行列式的特征值.

幸好，这些分子运动内在的物理规律为我们提供了一个较为简捷的途径. 我们不要先计算特征值后求特征向量,而是反过来,先求特征向量. 正如下文所述，这样就不会有太大的困难了. 当然,知道特征向量以后,它所对应的特征值与振动频率也立刻可以求出.

(i) 首先，该分子可能的纵向运动之一是以常速率作平移. 在这样的运动中,$x_1=x_2=x_3=x$,而方程是(参看(2)式)

$$\begin{pmatrix}\ddot{x}\\ \ddot{x}\\ \ddot{x}\end{pmatrix}=A\begin{pmatrix}x\\ x\\ x\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}=0\begin{pmatrix}x\\ x\\ x\end{pmatrix}.$$

因而，我们有了第一类正规方式：对应特征值 0 的平移.

(ii) 该系统关于碳原子对称,我们来找一类对称的振动方式,亦即 $x_1=-x_3, x_2=0$ 的振动方式.

我们也可以用下面的论证来说明这样做是对的. 因为两个氧原子毫无区别，并且在平衡状态下它们与碳原子的距离也是一样的. 所以,如果把 x_1 与 x_3 全部互换,运动方程(1)不受影响.（分子的对称性也反映在矩阵 A 中: $a_{12}=a_{32}, a_{13}=a_{31}, a_{21}=a_{23}, a_{11}=a_{33}$,也就是说,与下标 1 有关的矩阵元素总是与带下标 3 的对应元素相等.）

假如$\begin{pmatrix}a\\ b\\ c\end{pmatrix}$是一个特征向量(代表一个正规方式)，对应特征值 λ. 则把它的第一分量与第三分量交换以后,得到的向量仍然是对

应 λ 的特征向量．这两个向量的差(除非 $a=c$)

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} - \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} a-c \\ 0 \\ -(a-c) \end{pmatrix}$$

也是对应 λ 的特征向量，它的分量满足 $x_1=-x_3, x_2=0$。

(iii) 在上述向量中置 $a=c$，亦即 $x_1=x_3$，我们就得到第三类正规方式．又因为当此分子在平衡位置附近振动时，在每一状态下，该系统的线性动量都是零．所以

$$m\dot{x}_1+M\dot{x}_2+m\dot{x}_3=0.$$

由此得出

$$x_2=-\left(\frac{2m}{M}\right)x_1.$$

现在不难验证，上面所得到的 x_1, x_2 与 x_3 之间的每组关系式都定义了 A 的特征向量．对应的特征值在计算过程中也会自然地得到．

注意分子的正规振动方式与分子的对称性是相容的．在本例中，仅有一个非平凡的对称——一个反射，反射平面通过碳原子且垂直于两个氧原子所在的直线．振动着的分子在这样对称变换下的像与原来的振子完全一样．

三维空间中的正规方式

二氧化碳分子的一般运动是由九个坐标刻划的，每个原子三个坐标*(图 89)．在任一特定时刻，分子的位置(相对于平衡位置而言)用一个列向量 $\mathbf{v}$ 来表示，$\mathbf{v}$ 的九个分量是 $x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3$．这个位置向量的九维空间称为该分子的构形空间．运动方程可以写成

$$\ddot{\mathbf{v}}=B\mathbf{v} \tag{9}$$

* 这里我们放弃了把 y_i 用来表示正规坐标的做法，因为我们需要用这些记号参与表示该原子的特定坐标。

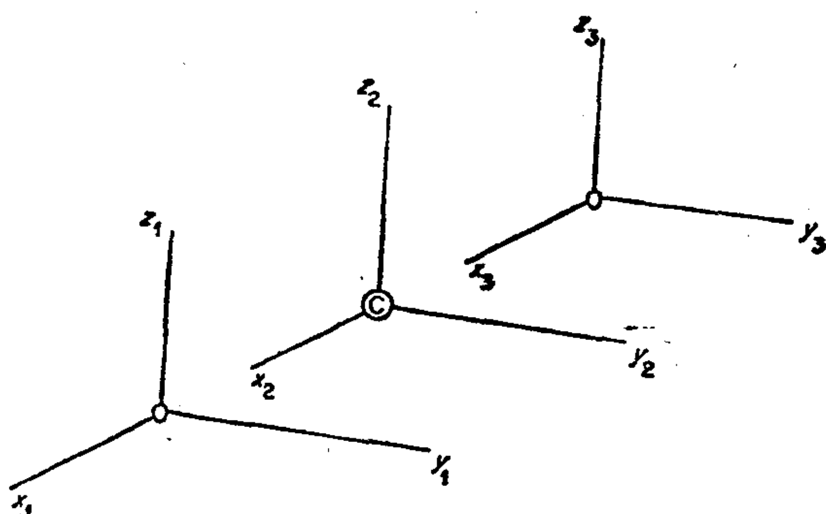


图 89 CO_2分子的坐标系

的形式，这里 B 是 9×9 矩阵．由此可见，最多可能有九类正规方式．我们已经找到其中三类，现在再来说明确实还有六类．

首先，在 y 方向与 z 方向上各有一个平移．特征向量由

(iv) $y_1=y_2=y_3, x_i=z_i=0$（对所有 i）；

(v) $z_1=z_2=z_3, x_i=y_i=0$（对所有 i）

给出．这两类正规方式的特征值都是 $\lambda=0$．

此外，还有两个一维的特征向量空间也对应原子的非加速运动，所以也是对应特征值 $\lambda=0$ 的．这些特征向量代表旋转（参看表 17 的图示）：

(vi) $y_1=-y_3, y_2=0, x_i=z_i=0$（对所有 i）；

(vii) $z_1=-z_3, z_2=0, x_i=y_i=0$（对所有 i）．

最后，还有两类正规振动方式，它们由下列特征向量给出：

(viii) $y_1=y_3, y_2=-\left(\dfrac{2m}{M}\right)y_1, x_i=z_i=0$（对所有 i）；

(ix) $z_1=z_3, z_2=-\left(\dfrac{2m}{M}\right)z_1, x_i=y_i=0$（对所有 i）．

其中碳原子分量与氧原子分量之间的关系仍由线性动量守恒推

出．这两个正规振动方式与分子的对称性也是相容的．

我们把所得到的结果总结在表 17 上．

表 17 CO_2 分子的正规方式

类型	三个平移			两个旋转		四个振动			
方式									
特征向量									
x_1	1	0	0	0	0	1	1	0	0
y_1	0	1	0	1	0	0	0	1	0
z_1	0	0	1	0	1	0	0	0	1
x_2	1	0	0	0	0	0	$-\frac{2m}{M}$	0	0
y_2	0	1	0	0	0	0	0	$-\frac{2m}{M}$	0
z_2	0	0	1	0	0	0	0	0	$-\frac{2m}{M}$
x_3	1	0	0	0	0	-1	1	0	0
y_3	0	1	0	-1	0	0	0	1	0
z_3	0	0	1	0	-1	0	0	0	1
参考课文中的条目	(i)	(iv)	(v)	(vi)	(vii)	(ii)	(iii)	(viii)	(ix)

用表中所列的向量作列的 9×9 矩阵 C 是非奇异的，并且 $C^{-1}BC$ 是对角矩阵．坐标变换 $\mathbf{v}=C\mathbf{v}^*$ 把方程组(9) 化简成用正规坐标表出的方程组，对后一方程组我们可以用观察法求解．在表 17 中，除了(viii)、(ix)两项中的特征频率未求出外，我们已经完成了这个问题的解答．

三维分子

三维分子振动问题的解法仅与 CO_2 分子的情况稍有差别．如果此分子有 n 个原子，则它在任一瞬间的位置由一个 $3n$ 维向量描述．它的运动方程涉及一个 $3n\times 3n$ 矩阵，它的非零特征值与

振动的固有频率以简单方式相联系．要求它的特征值，最好先找出它所对应的特征向量．而这些特征向量刻划了分子运动的正规方式．

在所有可能的 $3n$ 类正规方式中，三类是平移，三类是旋转（在线性分子的情况只有两类旋转），剩下的 $3n-6$（或 $3n-5$）类正规方式是振动．在这里我们不能仔细分析如何求出这些正规方式，但想指出，这些正规振动方式与分子的对称性密切相关．在 CO_2 分子中，每一正规振动方式在对称变换下不变．对于更一般的分子，这一点不再成立．但是，这些正规方式可以分成若干组，使得该分子的对称变换群总是把一组中的正规方式映到本组正规方式的线性组合．因此，运用该分子对称变换群结构的知识，不难把这些正规方式组以及它们所生成的向量空间求出．这样，我们获得群论在物理学中的一个重要应用．

分子振动的动力学理论在下面的参考书中论及：

Corben & Stehle, *Classical Mechanics*, Second Edition, Wiley, 1960.

有关群论的应用，可参阅下列书籍：

1. F. A. Cotton, *Chemical Applications of Group Theory*, Wiley, 1962.
2. R. McWeeny, *Symmetry*, Pargamon, 1963.
3. Jaffe & Orchin, *Symmetry in Chemistry*, Wiley, 1965.
4. L. F. Phillips, *Basic Quantum Chemistry*, Wiley, 1965.
5. R. M. Hochstrasser, *Molecular Aspects of Symmetry*, W. A. Benjamin, 1966.
6. G. G. Hall, *Applied Group Theory*, Longmans, 1967.

练　习

1．水分子 H_2O(图 67)不是线性的．在平衡状态下，两个氢原子与氧原子保持同样的距离，但它们不在一条直线上．所以，在九类可能的正规方式中，三类是平移，三类是旋转．试画出三类线性无关的正规振动方式的草图．

［提示：这些正规振动方式与分子的对称性相容．］

18.2　相对论：洛伦兹变换

数学中最重要的概念之一是不变性的概念．这种思想对我们来说已不是新的了．例如，我们知道了一个物体的对称性是用保持该物体不变的等距变换来衡量的．又如，在上一节中我们提到过一个系统的运动方程在对称变换下是不变的，并且指出了怎样才能以这种思想为指导，去求得正规振动方式．

在狭义相对论中我们还可以举出另外一个例子．狭义相对论的基本假设如下：

(i) 对于两个以常速度相对运动的观察者来说，每一个物理定律都同等地有效．

(ii) 光速是一个物理常数 c：两个以常速度相对运动的观察者都会观察到光在真空中沿直线以速率 c 运动．(实验证明，c 近似于 300000 公里/秒．)

考虑两个以常速度相对运动的观察者．每一观察者设置一个直角坐标系．假定在一个被两个观察者一致称为 $t=0$ 的时刻，这两个坐标系瞬时地重合．然后坐标系随着各自的观察者运动．

第一个观察者注意到，在时间 t 内，光从 $(0,0,0)$ 运动到了 (x,y,z)．他得出光速 c 是由 $(x^2+y^2+z^2)^{\frac{1}{2}}=ct$ 或者

$$x^2+y^2+z^2-c^2t^2=0 \qquad (10)$$

给出的．

第二个观察者注意到，在他的坐标系中，光在时间 t' 中由(0，

0,0)运动到了(x',y',z')．他得出光速 c 由下式给出：

$$x'^2+y'^2+z'^2-c^2t'^2=0. \tag{11}$$

由基本假设，在方程(10)与(11)中的常数 c 相同．这两个方程给出了两个坐标系间的一个关系：

$$x^2+y^2+z^2-c^2t^2=x'^2+y'^2+z'^2-c^2t'^2. \tag{12}$$

设 $T:(x,y,z,t)\longrightarrow(x',y',z',t')$ 是联系第一个观察者的量值与第二个观察者的对应量值的变换，则二次型

$$Q=x^2+y^2+z^2-c^2t^2 \tag{13}$$

在变换 T 下不变．

因而，出现一个有趣的问题：我们能够找到一个变换 T 保持 Q 不变吗？特别，T 能够取为线性变换吗？

我们只限于注意这种情况：第二个观察者以速率 v 在 x 轴方向上相对于第一个观察者运动（图 90）．此时坐标平面 $y=0$ 与 $y'=0$ 永远重合；平面 $z=0$ 与 $z'=0$ 也是这样．于是，我们总可以假定 $y=y'$ 及 $z=z'$．条件(12)就简化为

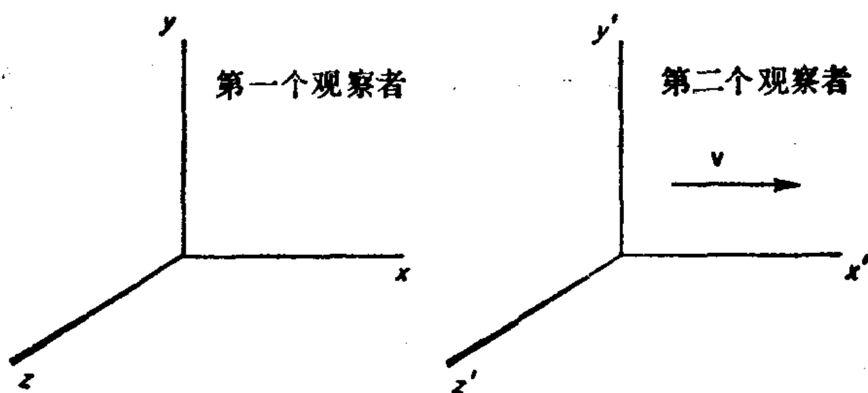


图 90　两个观察者的参考系

$$x^2-c^2t^2=x'^2-c^2t'^2. \tag{14}$$

设 T 是一个线性变换，它把(x,t)映到(x',t')，并且使条件(14)满足．首先注意到 $x'=0$ 蕴涵着 $x-vt=0$，因为第二个观察者的坐标原点 O′ 以速率 v 相对于第一个观察者运动．由此可见，对于某

个(可能与 v 有关的)常数 γ，有

$$x'=\gamma(x-vt). \tag{15}$$

根据坐标系的取法(图 90)，可以假定 γ 是正的(考虑当 $t=0$ 时的量值). 我们再用线性关系

$$t'=ax+bt \tag{16}$$

完成 T 的定义，这里 a 与 b 是待定的常数.

把(15)、(16)两式代入(14)式，我们得知二次型 $x^2-c^2t^2$ 恒等于

$$\gamma^2(x-vt)^2-c^2(ax+bt)^2.$$

比较 x^2、xt 以及 t^2 项的系数，我们得到

$$1=\gamma^2-a^2c^2, \tag{17}$$

$$0=-\gamma^2v-abc^2, \tag{18}$$

$$-c^2=\gamma^2v^2-b^2c^2. \tag{19}$$

由方程(18)

$$a^2b^2c^4=\gamma^4v^2; \tag{20}$$

但由方程(17)与(19)

$$a^2b^2c^4=(\gamma^2-1)(\gamma^2v^2+c^2). \tag{21}$$

最后，由方程(20)与(21)，我们求出了正常数

$$\gamma=\gamma(v)=\frac{1}{\sqrt{1-v^2/c^2}}. \tag{22}$$

再代入到方程(17)与(19)中去，我们得到

$$a=a(v)=\pm\frac{v/c^2}{\sqrt{1-v^2/c^2}},$$

$$b=b(v)=\pm\frac{1}{\sqrt{1-v^2/c^2}}. \tag{23}$$

注意 γ, a, b 都是与 v 有关的常数，并且

$$\gamma(-v)=\gamma(v),\quad a(-v)=-a(v),\quad b(-v)=b(v).$$

为了保证这些常数是实的，必须$|v|<c$.

只剩下要确定a与b的符号了．线性变换T若用矩阵形式给出是

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = L(v)\begin{pmatrix} x \\ t \end{pmatrix}, \quad 其中 L(v)=\begin{pmatrix} \gamma & -\gamma v \\ a & b \end{pmatrix} \tag{24}$$

（参看(15)式与(16)式）．对第一个观察者而言，一个被他用(x,t)度量的事件被第二个观察者用(x',t')度量，坐标通过(24)式相联系．而对第二个观察者而言，第一个观察者以速度$-v$运动，所以，他将按照

$$\begin{pmatrix} x \\ t \end{pmatrix} = L(-v)\begin{pmatrix} x' \\ t' \end{pmatrix} \tag{25}$$

来计算量值之间的关系．从(24)式与(25)式，我们得到

$$L(v)L(-v)=I. \tag{26}$$

因此

$$\begin{pmatrix} \gamma & -\gamma v \\ a & b \end{pmatrix}\begin{pmatrix} \gamma & \gamma v \\ -a & b \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

容易求得

$$b=\gamma=\frac{1}{\sqrt{1-v^2/c^2}}, \quad a=\frac{-v/c^2}{\sqrt{1-v^2/c^2}}. \tag{27}$$

线性变换

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = L(v)\begin{pmatrix} x \\ t \end{pmatrix}, \quad L(v)=\gamma(v)\begin{pmatrix} 1 & -v \\ -v/c^2 & 1 \end{pmatrix}, \quad |v|<c \tag{28}$$

称为洛伦兹变换．它们是联系以常速度v相对运动的观察者的量值的线性变换，并且满足相对论的基本假设．特别，$x^2-c^2t^2$在洛伦兹变换下不变．

矩阵的集合$\{L(v)$，对所有使$|v|<c$的$v\}$在矩阵乘法下成为

一个群，称为洛伦兹群．由(26)式显然有

$$[L(v)]^{-1}=L(-v)。$$

此外，可以证明

$$L(v)L(v')=L(v''),$$

其中

$$v''=\frac{v+v'}{1+\frac{vv'}{c^2}}. \tag{29}$$

这就是狭义相对论中速度迭加的法则．

最后，我们希望指出上文的讨论同偏微分方程的联系．微分方程

$$\frac{\partial^2 U}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2}=0 \tag{30}$$

在洛伦兹变换下不变，亦即，根据变换(28)把变量 x,t 变为 x',t' 后，微分方程(30)成为

$$\frac{\partial^2 U'}{\partial x'^2}-\frac{1}{c^2}\frac{\partial^2 U'}{\partial t'^2}=0, \tag{31}$$

这里 U' 是 x' 与 t' 的函数，它是由 U 经过变量变换而得到的．注意方程(31)与方程(30)是相同的方程，只不过后者用新变量表出．当我们把方程(30)看成以速度 c 运动的波(例如光波)的波动方程时，这种不变性的理由是显然的．因为根据狭义相对论，波所满足的微分方程在两个观察者的坐标系中具有相同的形式．

如欲深入学习，建议阅读如下书籍：

1. W. Rindler, *Special Relativity*, Oliver & Boyd, 1960.

2. J. L. Synge, *Relativity; the Special Theory*, Wiley, 1968.

3. Corben & Stehle, *Classical Mechanics*, Second Edition. Wiley, 1960，书中专述狭义相对论的一章．

练　　习

2．证明速度迭加的法则(29)。

3．观察者 A 以速度 $\frac{3}{4}c$ 相对于观察者 B 运动，而观察者 B 在同一方向上以速度 $\frac{3}{4}c$ 相对于观察者 C 运动．求 A 相对于 C 的速度。

4．验证，矩阵 $L(v)$ 的行列式值为 1。

5．验证 $[L(v)]^{-1}=L(-v)$。

6．证明，所有由(28)式定义的矩阵 $L(v)$ 的集合在矩阵乘法下成为一个群。

7．(i) 验证，函数 $U(x,t)=\sin(x-ct)$ 满足波动方程(30)；

(ii) 求使用洛伦兹变换(28)后所得到的函数 $U'(x',t')$。

[提示：

$$\begin{pmatrix} x \\ t \end{pmatrix}=[L(v)]^{-1}\begin{pmatrix} x' \\ t' \end{pmatrix}=L(-v)\begin{pmatrix} x' \\ t' \end{pmatrix}.$$

因而

$$x=\gamma x'+\gamma vt',\quad t=\left(\frac{\gamma v}{c^2}\right)x'+\gamma t'.\ ]$$

(iii) 验证，函数 U' 也满足波动方程。

8．按如下方法一般地证明波动方程在洛伦兹变换下不变：设两组变量 x,t 与 x',t' 通过变换(28)相联系，并且 U' 是 x',t' 的函数，它是把函数 $U(x,t)$ 中的变量 x,t 用 x',t' 表出后得到的．证明

$$\frac{\partial^2 U}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2}=\frac{\partial^2 U'}{\partial x'^2}-\frac{1}{c^2}\frac{\partial^2 U'}{\partial t'^2}.$$

18.3　概率论：马尔可夫过程

在物理学、生物学与社会科学中有一大类问题，它们研究的是事件的一个序列，其中每一事件仅与紧挨在它前面的事件有关．这样一个事件序列称为马尔可夫过程，以纪念俄罗斯数学家 A.A.

马尔可夫，他在 1907 年首先提出了这个论题．本节中我们要举几个例子，并且简单地讨论一下遇到的代数问题．

例 1 人口研究

某国对城镇与农村之间的人口流动作年度调查，发现有一个稳定的朝城镇流动的趋势，具体情况如下：

(i) 每年，农村居民的 $2\frac{1}{2}\%$ 移居城镇；

(ii) 每年，城镇居民的 1% 移居农村．

现在，总人口的 60% 住在城镇．假定总人口（包括城镇与农村）保持不变，并且人口流动的这种趋势继续下去．那么，一年以后住在城镇的人口所占的比例是多少？二年以后呢？十年以后呢？最终呢？

看起来好象城镇居民在总人口中所占的比例将逐年增加．但是，如果由此推断最终全国人口都要住在城镇，那就非常荒谬了．

这个问题适于用矩阵方法解决．设 t_0 与 c_0 分别为现在城镇与农村居民所占的比例，又设 t_k 与 c_k 为 k 年以后对应的比例．在我们的问题中，$t_0=0.6, c_0=0.4$．我们先求 t_1 与 c_1．假定总人口为 N．据假设，这个数在任何时候都不变．于是，一年以后，城镇人口 t_1N 是由原来城镇居民的 99% 加上原来农村人口的 $2\frac{1}{2}\%$ 而组成的，亦即

$$t_1N=0.99\,t_0N+0.025\,c_0N. \tag{32}$$

同理

$$c_1N=0.01\,t_0N+0.975\,c_0N. \tag{33}$$

这样，我们问题的第一部分就解决了：置 $t_0=0.6, c_0=0.4$，我们得到 $t_1=0.604, c_1=0.396$．所以，一年以后，总人口的 60.4% 住在城镇．

方程(32)与(33) 给出了 t_1, c_1 与 t_0, c_0 之间的线性关系．这两个方程可以用矩阵形式写成

$$(t_1 \quad c_1)=(t_0 \quad c_0)\begin{pmatrix}0.99 & 0.01\\ 0.025 & 0.975\end{pmatrix}. \qquad (34)$$

我们把关系式用行向量写出．这样就与大部分有关这一问题的文献中的写法一致．矩阵

$$P=\begin{pmatrix}0.99 & 0.01\\ 0.025 & 0.975\end{pmatrix} \qquad (35)$$

描述了从现在到一年之后的转变．又因为假定人口流动的这一趋势持续下去，所以这一矩阵同样描述了 k 年之后到 $k+1$ 年之后的转变：

$$(t_{k+1} \quad c_{k+1})=(t_k \quad c_k)P. \qquad (36)$$

从(36)式显然可以看出，我们在同马尔可夫过程打交道：$k+1$ 年后的情况仅仅同紧挨在它前面的那一年(第 k 年)的情况有关．矩阵 P 称为这一马尔可夫过程的转移矩阵．在(36)式中置 $k=1$，并且用(34)式，我们有

$$(t_2 \quad c_2)=(t_1 \quad c_1)P=(t_0 \quad c_0)P^2.$$

同理可得，对所有的 k

$$(t_k \quad c_k)=(t_0 \quad c_0)P^k. \qquad (37)$$

因此，矩阵 P^k描述了从现在到 k 年之后的转变．特别

$$P^2=\begin{pmatrix}0.98035 & 0.01965\\ 0.049125 & 0.950875\end{pmatrix},$$

由此，$t_2=0.98035\,t_0+0.049125\,c_0=0.60786$．因而，二年以后，总人口的 60.786%是城镇居民．

十年以后的人口分布由矩阵 P^{10}所支配，而最终的人口分布则是由当 k 充分大时，P^k的性质决定的．我们先找一个计算 P^k 的公式，然后再用这个公式把剩下的两个问题一块解决．首先，我们再仔细考察一下矩阵 P．如果我们把 P 的行与列分别标上 t（城镇）与 c（农村），则它描述从一年到下一年转变的方法就看得

很清楚了：

$$\begin{matrix} & t & c \\ t & 0.99 & 0.01 \\ c & 0.025 & 0.975 \end{matrix}$$

$$\begin{pmatrix} 0.99 & 0.01 \\ 0.025 & 0.975 \end{pmatrix}.$$

如果一个人住在城镇，我们就说他处于状态 t；如果他住在农村，就说他处于状态 c. 那么，在标以 i 的行与标以 j 的列交叉处的矩阵元素表示原处于状态 i 而一年后处于状态 j 的人口在原处于状态 i 的人口中所占的比例. 例如，ct-元素($=0.025$)给出了在一年期间要迁居到城镇的农村居民在农村居民中所占的比例.

因为每个城镇居民不是留在城镇，就是迁到农村，所以 P 的第一行元素的和是 1. 同样的结论对 P 的第二行也成立. 因而，P 是下面定理所考虑的那一类型的矩阵.

定理 I 设

$$P=\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad 且\ p+q\neq 0,$$

则 P^k 由公式

$$(p+q)P^k=\begin{pmatrix} q & p \\ q & p \end{pmatrix}+r^k\begin{pmatrix} p & -p \\ -q & q \end{pmatrix} \tag{38}$$

给出，其中 $r=1-(p+q)$.

证明 对 k 用归纳法，不难证明公式(38). 但这个证法并不能阐明最初是怎样得到这个公式的. 我们介绍一个证明方法，不需要预先知道 P^k(对任意的 k)的表达式.

我们的证明基于如下事实：

(i) 对角矩阵的 k 次幂可以即刻写出；

(ii) P 相似于对角矩阵 D：存在一个矩阵 C，使 $C^{-1}PC=D$. 此时

$$D^k=(C^{-1}PC)^k=C^{-1}P^kC,$$

所以

$$P^k = CD^kC^{-1}. \tag{39}$$

我们着手对角化P. 它的特征方程是

$$|P-\lambda I| = (\lambda-1)(\lambda-1+p+q) = 0,$$

所以它的特征值为1与$1-(p+q)$. 置$r=1-(p+q)$. 特征向量空间的生成元如下所示：

$$\mathscr{S}_1:\ \begin{pmatrix}1\\1\end{pmatrix};\quad \mathscr{S}_r:\ \begin{pmatrix}-p\\q\end{pmatrix}.$$

据13.3节的定理V,矩阵

$$C=\begin{pmatrix}1 & -p\\1 & q\end{pmatrix}\quad 使\quad C^{-1}PC=\begin{pmatrix}1 & 0\\0 & r\end{pmatrix}.$$

所以

$$P^k = C\begin{pmatrix}1 & 0\\0 & r\end{pmatrix}^k C^{-1} = C\begin{pmatrix}1 & 0\\0 & r^k\end{pmatrix}C^{-1}, \tag{40}$$

据此，公式(38)即得. 因为$|C|=p+q$，条件$p+q\neq 0$保证了C是非奇异的.

在我们的人口问题中

$$p=0.01,\ q=0.025,\ r=1-(p+q)=0.965.$$

现在只要把这些值以及$k=10$代入(38)式，便可求得十年后的人口分布了.

最终人口分布情况如何呢？因为$0<r<1$，当$k\to\infty$时，$r^k\to 0$. 所以,从公式(38)得到

$$当\ k\to\infty 时，P^k \to \frac{1}{p+q}\begin{pmatrix}q & p\\q & p\end{pmatrix}. \tag{41}$$

若把城镇与农村的居民最终所占的比例用行向量(t_∞, c_∞)表示，则由(37)式知

$$(t_\infty \quad c_\infty)=(t_0 \quad c_0)\begin{pmatrix}\frac{q}{p+q} & \frac{p}{p+q}\\ \frac{q}{p+q} & \frac{p}{p+q}\end{pmatrix}$$

$$=\left(\frac{q(t_0+c_0)}{p+q} \quad \frac{p(t_0+c_0)}{p+q}\right).$$

但 $t_0+c_0=1$. 从而

$$(t_\infty \quad c_\infty)=\left(\frac{q}{p+q} \ \frac{p}{p+q}\right). \tag{42}$$

在我们的问题中，$p=\frac{1}{100}$, $q=\frac{1}{40}$. 所以，最终，总人口的$\frac{5}{7}(=t_\infty)$住在城镇.

注意这一结果与 t_0、c_0 无关. 所以，不管最初的人口分布如何，最终城镇和农村的居民是按 5:2 的比率分布的. 这个最终分布是城镇和农村之间的平衡状态：迁到农村的城镇居民正好被迁到城镇的农村居民所抵消.

例 2 在扩散理论中的应用

假定某一物质能以液态与气态存在. 又设在任意一段很短的时间内

(i) 液体的 $2\frac{1}{2}\%$ 蒸发；

(ii) 气体的 1% 凝结.

那么，不管最初的液-气比率如何，最终将达到一个平衡状态，此时物质的$\frac{5}{7}$是气态的.

其分析过程同例 1 一样，只不过以“液体”代替“农村”，以“气体”代替“城市”. 实际上，液体气体的互相变化是一个连续过程，我们的离散过程只能作为一个简化的模型.

例 3 随机游动问题

一类很重要的马尔可夫过程是所谓“随机游动问题”. 这类问

题有其广泛的应用，例如在扩散与布朗运动的理论中，在赌博中也是如此．考虑下面的简单例子：

一个醉汉在他的床（图 91 上标为 4）与离床三步远的楼梯顶（图上标为 1）之间蹒跚．他每走一步，朝着床走的机会与朝着楼梯走的机会的比是 3 比 1．如果他到了位置 1，就会摔下楼梯；如果他到了位置 4，就一定会睡个好觉．证明，这个醉汉最终*一定到达床和楼梯顶这两个位置中的一个．假定他最初

(i) 离床二步远(在位置 2)；

(ii) 离床一步远(在位置 3)，

计算他到床上去的概率．

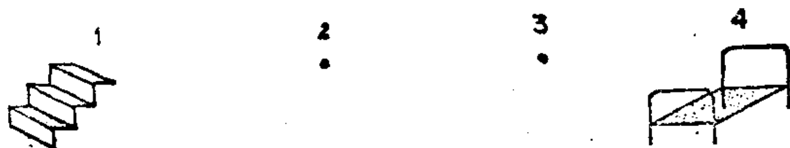


图 91　醉汉走路

我们来作一个 4×4 转移矩阵 $P=(p_{ij})$，这里 p_{ij} 表示当醉汉在位置 i 时，下一步他走到位置 j 的可能性．因此，$p_{44}=1$，而 $p_{4j}=0$ $(j=1,2,3)$，因为他一到位置 4，立刻就躺到床上去了．位置 4（床）称为吸收态，因为到了这个位置就不可能再离开了．位置 1 也是吸收态：$p_{11}=1$，而 $p_{1j}=0$ $(j=2,3,4)$．

当醉汉在位置 2 或 3 时，他朝着床走一步的概率是 $\frac{3}{4}$，背着床走一步的概率是 $\frac{1}{4}$．因此，$p_{23}=\frac{3}{4}$，$p_{21}=\frac{1}{4}$，$p_{22}=0$（他停留在 2 的机会不存在），$p_{24}=0$（他不能一步从 2 跨到 4）．同理，$p_{31}=p_{33}=0$，$p_{32}=\frac{1}{4}$，$p_{34}=\frac{3}{4}$．整个矩阵是

* 假定他在走完床与楼梯顶之间这一段路之前不会清醒过来，也不会摔倒。

$$P=\begin{pmatrix}1 & 0 & 0 & 0\\ \frac{1}{4} & 0 & \frac{3}{4} & 0\\ 0 & \frac{1}{4} & 0 & \frac{3}{4}\\ 0 & 0 & 0 & 1\end{pmatrix}. \tag{43}$$

注意 P 的所有元素是非负的(因为都是概率)，并且每一行元素的和是 1 . 具有这两个性质的正方矩阵称为随机矩阵. 例 1 的矩阵(35)也是随机矩阵. 现在假定醉汉在位置 2 . 这个事实可以用行向量 (0 1 0 0) 表示. 这个向量可以解释为醉汉在位置 1, 3, 4 的概率都是零，而他一定在位置 2 . 走一步以后，他的位置概率是

$$(0\ 1\ 0\ 0)P=\left(\frac{1}{4}\ 0\ \frac{3}{4}\ 0\right).$$

假如他再走一步，那么，根据概率论的法则，他的位置概率是

$$\left(\frac{1}{4}\ 0\ \frac{3}{4}\ 0\right)P=\left(\frac{1}{4}\ \frac{3}{16}\ 0\ \frac{9}{16}\right). \tag{44}$$

也就是说，走两步以后，他在位置 1, 2, 4 的概率分别为 $\frac{1}{4}$, $\frac{3}{16}$, $\frac{9}{16}$；他肯定不在位置 3 . 注意

$$\left(\frac{1}{4}\ 0\ \frac{3}{4}\ 0\right)P=(0\ 1\ 0\ 0)P^2,$$

所以，这个醉汉走两步的运动受矩阵 P^2 支配. 更一般地，P^k 是他走 k 步这一运动的转移矩阵.

我们感兴趣的是醉汉的最终命运. 用代数语言说，我们希望知道当 $k\to\infty$ 时，P^k 的情况究竟如何. 方法之一是像例 1 那样，先把 P 对角化(如果可能)，求出计算 P^k 的公式，然后令 $k\to\infty$，再说明 P^k 趋近于一个极限矩阵 L. 这个方法牵涉到一些冗长的

计算，而如果我们采用下面的方法，这些计算是可以避免的.

我们首先证明这个醉汉最终总会到达床或楼梯顶. 因若不然，他必须一直在位置 2 与 3 之间来回走动. 假定他从 2 出发，则他走到 3 再回到 2 这样来回 k 次的概率为

$$\left(\frac{3}{4}\cdot\frac{1}{4}\right)^k=\left(\frac{3}{16}\right)^k.$$

从 3 出发来回 k 次的概率也是一样. 而当 $k\to\infty$ 时，$\left(\frac{3}{16}\right)^k\to 0$. 因此，他一直这样来回走动的概率是零. 由此可见，这个醉汉一定会在位置 1 或 4 结束这一过程.

假定醉汉从位置 2 出发，最终到达楼梯顶的概率为 p；如果他从位置 3 出发，此概率为 q. 那么，相应的到达床的概率分别为 $1-p$ 与 $1-q$. 因而，醉汉的最终命运是由矩阵

$$L=\begin{pmatrix}1&0&0&0\\p&0&0&1-p\\q&0&0&1-q\\0&0&0&1\end{pmatrix}\tag{45}$$

描述的. 我们再从关系式

$$PL=L\tag{46}$$

求 p 与 q. 这个关系式是成立的，因为

$$L=\lim_{k\to\infty}P^k=\lim_{k\to\infty}P^{k+1}=P(\lim_{k\to\infty}P^k)=PL.$$

从(46)式推出

$$p=\frac{1}{4}+\frac{3q}{4},$$

$$q=\frac{p}{4}.$$

所以 $q=\frac{1}{13}$，$p=\frac{4}{13}$. 这个醉汉从位置 2 出发，到达床的概率为

$\frac{9}{13}$；从位置 3 出发，此概率为$\frac{12}{13}$. 本题的解答至此完成了.

马尔可夫过程在物理学、生物学和社会科学的许多分支中都起重要的作用. 例如，它是遗传理论与所谓“生灭过程”理论的基础. 我们建议把下列书籍作为参考书：

1. W. Feller, *An Introduction to Probability Theory*, Volume 1, Third Edition, Wiley, 1968.

2. Kemeny, Snell & Thompson, *Introduction to Finite Mathematics*, Prentice-Hall, 1964.

3. S. Lipschutz, *Theory and Problems of Finite Mathematics*, Schaum, 1966.

4. Mathews & Langenhop, *Discrete and Continuous Methods in Applied Mathematics*, Wiley, 1966.

5. S. R. Searle, *Matrix Algebra for the Biological Sciences*, Wiley, 1966.

练　　习

9. 设 P 与 Q 是 $n\times n$ 随机矩阵. 证明，乘积 PQ 与 $kP+(1-k)Q$ $(0\leqslant k\leqslant 1)$ 都是随机矩阵.

10. 证明，所有随机矩阵都有特征值 1，并且求出该特征值对应的一个特征向量.

11. 一个三态马尔可夫过程的转移矩阵是

$$P=\begin{pmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{4}\\ \frac{3}{4} & 0 & \frac{1}{4}\\ 0 & 0 & 1\end{pmatrix}.$$

(i) 如果该系统从状态 2 开始，求经过两个步骤以后又回到状态 2 的概率.

(ii) 证明，状态 3 是吸收态.

(iii) 证明，该系统最终一定吸收到状态 3。最终不可能回到的状态称为瞬态。因此，状态 1 与 2 都是瞬态。

12. 若 P 如练习 11 所示，求计算 P^k 的公式。

13. 所有元素都是正数的矩阵称为正矩阵。给定

$$P=\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \end{pmatrix},$$

证明 P^3 是正矩阵。

14. 设 $P=(p_{ij})$ 与 $Q=(q_{ij})$ 是 3×3 随机矩阵，其中 P 是正矩阵。置 $PQ=R=(r_{ij})$。证明

$$r_{ij}\geqslant q_{1j}, q_{2j}, q_{3j} \text{中最小者},$$

等号成立当且仅当 $q_{1j}=q_{2j}=q_{3j}$。

[提示：用 $p_{i1}+p_{i2}+p_{i3}=1$ 这一事实；此外，在最后一步再用 $p_{ij}>0$（对所有 i,j）。]

15. 利用练习 14 证明，如果 P 是 3×3 正的随机矩阵，则 $\lim\limits_{k\to\infty}P^k$ 一定存在，并且是形如

$$\begin{pmatrix} a & b & c \\ a & b & c \\ a & b & c \end{pmatrix}, \quad a+b+c=1$$

的正矩阵。

[提示：矩阵 $P, P^2, P^3, \cdots$ 的 ij-元素组成一个单调递增序列，并且有上界 1。这样的序列是收敛的。]

16. 设 P 是练习 13 所示的矩阵。证明

$$\lim_{k\to\infty} P^k=\frac{1}{25}\begin{pmatrix} 12 & 8 & 5 \\ 12 & 8 & 5 \\ 12 & 8 & 5 \end{pmatrix}.$$

17. 证明，转移矩阵为

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \end{pmatrix}$$

的马尔可夫过程中没有任何一个状态是吸收态或瞬态。这些状态称为循环态．证明，这三个状态循环出现的频率比接近于 12:8:5．

18．设P是随机矩阵

$$P=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

证明，$\lim_{k\to\infty} P^k$ 不存在．注意，P 的任何次幂都不是正矩阵．

18.4 四端网络：传递矩阵；滤波器

本节中我们要举例说明矩阵论在电路理论中的应用．

考虑一个四端网络(图 92)，在此网络中稳定状态的输入和输出交变电动势与电流分别由

$$e_1(t)=E_1e^{j\omega t}, \qquad i_1(t)=I_1e^{j\omega t},$$

与

$$e_2(t)=E_2e^{j\omega t}, \qquad i_2(t)=I_2e^{j\omega t}$$

给出．这里，按照电路理论的习惯，j 代表复数$\sqrt{-1}$．物理量 E_1，

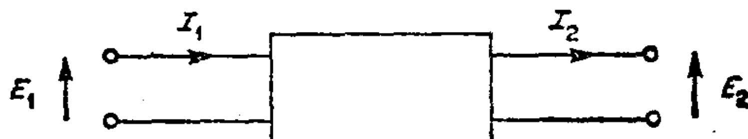


图 92　四端网络

I_1, E_2, I_2 为该系统相应的(复)振幅．

如果在一个网络中，输入电动势与电流可以表为输出电动势与电流的线性组合

$$\begin{aligned} e_1(t)&=ae_2(t)+bi_2(t), \\ i_1(t)&=ce_2(t)+di_2(t), \end{aligned} \tag{47}$$

则称此网络为线性网络，这里 a,b,c,d 为与时间 t 无关的复数，但是它们可能与频率 ω 有关．消去 $e^{j\omega t}$，我们得到矩阵方程

$$\begin{pmatrix} E_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} E_2 \\ I_2 \end{pmatrix}. \tag{48}$$

矩阵

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

称为这个四端网络的传递矩阵．

例 4　串联阻抗

由一个串联阻抗 Z 和一条理想导体的回路便可组成一个非常简单的四端网络(图 93)．我们回忆一下

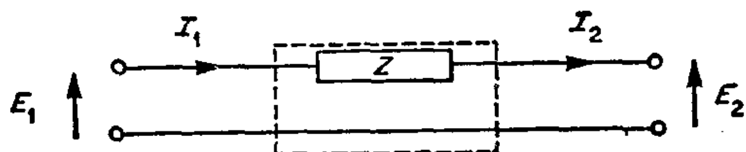


图 93　串联阻抗

$$Z = R + j\left(\omega L - \frac{1}{\omega C}\right),$$

其中 R，L 与 C 分别表示构成这一阻抗的电阻、电感与电容．由基尔霍夫定律，我们得到

$$E_1 = E_2 + Z I_2,$$
$$I_1 = \qquad I_2.$$

传递矩阵是

$$\begin{pmatrix} 1 & Z \\ 0 & 1 \end{pmatrix}. \tag{49}$$

注意，它的行列式值为 1．

例 5　分流阻抗

图 94 表示一个分流阻抗．由基尔霍夫定律，我们得到

$$E_1 = E_2,$$

$$I_1 = \left(\frac{1}{Z}\right) E_2 + I_2.$$

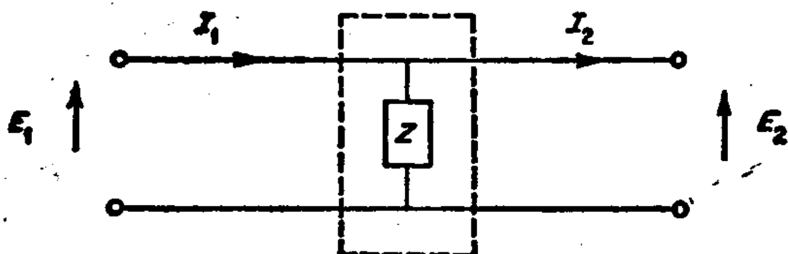


图 94 分流阻抗

传递矩阵是

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{Z} & 1 \end{pmatrix}. \tag{50}$$

这个矩阵的行列式值也是 1.

四端网络的级联

图 95 表示两个四端网络的级联，这两个网络的传递矩阵分别为 A 与 B. 我们有

$$\begin{pmatrix} E_1 \\ I_1 \end{pmatrix} = A \begin{pmatrix} E_2 \\ I_2 \end{pmatrix}, \quad \begin{pmatrix} E_2 \\ I_2 \end{pmatrix} = B \begin{pmatrix} E_3 \\ I_3 \end{pmatrix}.$$

所以

$$\begin{pmatrix} E_1 \\ I_1 \end{pmatrix} = AB \begin{pmatrix} E_3 \\ I_3 \end{pmatrix}.$$

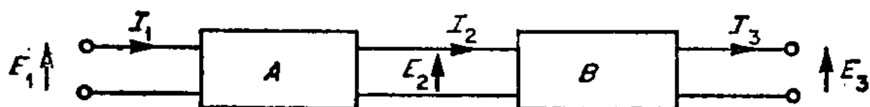


图 95 级联

因此，该系统与具有传递矩阵 AB 的单个四端网络等效. 不难把它推广成下面的结果:

定理 II 如果把传递矩阵分别为 $A_1, A_2, \cdots, A_n$ 的 n 个四端网络（依次）级联，所得到的四端网络的传递矩阵为 $A_1A_2\cdots A_n$.

例 6 求图 96 所示的"T 型"网络的传递矩阵.

该网络可以看成由三个网络级联而成的：一个串联阻抗 Z_1，接一个分流阻抗 Z_2，再接一个串联阻抗 Z_3. 由矩阵(49)与(50)，所得到的网络的传递矩阵为

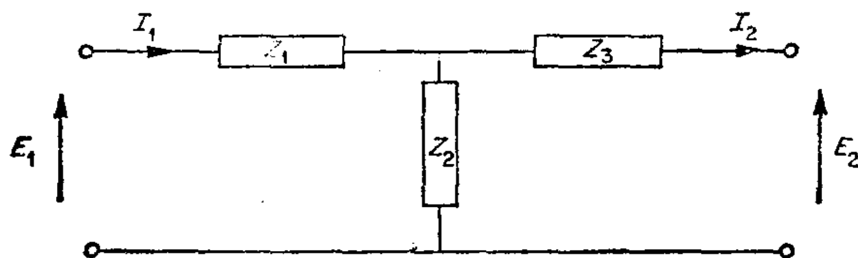


图 96 T 型网络

$$\begin{pmatrix} 1 & Z_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{1}{Z_2} & 0 \end{pmatrix}\begin{pmatrix} 1 & Z_3 \\ 0 & 1 \end{pmatrix}=\frac{1}{Z_2}\begin{pmatrix} Z_1+Z_2 & Z_1Z_2+Z_2Z_3+Z_3Z_1 \\ 1 & Z_2+Z_3 \end{pmatrix}. \quad (51)$$

根据乘积的行列式的定理，显然它的行列式值也是 1.

滤波器电路

一类很重要的网络是由大量相同的四端网络级联而成. 这样的网络称为滤波器，因为它只允许某些频率的电流自由通过，而把其他频率的电流减弱. 最普通的滤波器电路是所谓"梯级系统"(图 97). 图中 $Z_入$ 与 $Z_出$ 分别表示输入阻抗与输出阻抗. 整个电

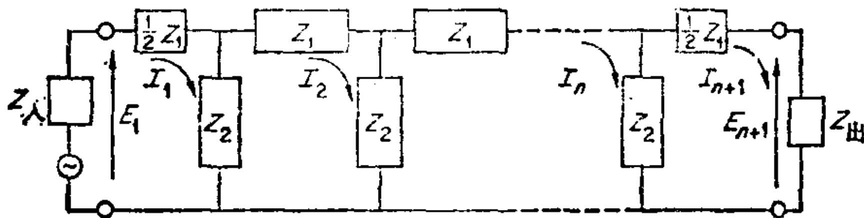


图 97 梯级系统

路可以看成 n 个对称的T型网络(图 98)级联而成.

假定角频率为 ω 的交变电流通过该电路. 每个T型网络的传递矩阵 A 是把 (51) 式的矩阵中的 Z_1 与 Z_3 都换成 $\frac{1}{2}Z_1$ 而得到的. 因此

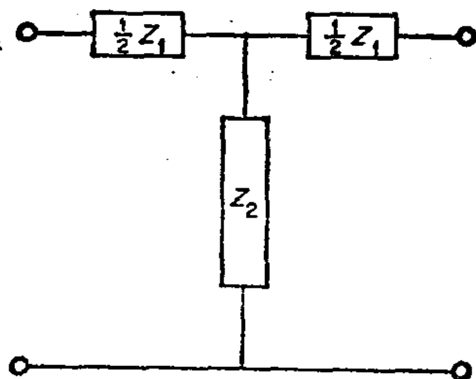


图 98　对称T型网络

$$A=\begin{pmatrix} a & b \\ c & a \end{pmatrix},$$

其中

$$a=\frac{\frac{1}{2}Z_1+Z_2}{Z_2},\quad b=\frac{\frac{1}{4}Z_1^2+Z_1Z_2}{Z_2},\quad c=\frac{1}{Z_2}. \tag{52}$$

我们必须注意, Z_1, Z_2 与 A 都是 ω 的函数.

矩阵 A 是矩阵(51)的特殊情况,我们有

$$|A|=a^2-bc=1. \tag{53}$$

据定理 II,输入与输出的振幅满足关系式

$$\begin{pmatrix} E_1 \\ I_1 \end{pmatrix}=A^n\begin{pmatrix} E_{n+1} \\ I_{n+1} \end{pmatrix}. \tag{54}$$

我们用把 A 对角化的方法去求计算 A^n 的公式. 用(53)式, 我们求出 A 的特征多项式为

$$|A-\lambda I|=\lambda^2-2a\lambda+1=(\lambda-\lambda_1)(\lambda-\lambda_2). \tag{55}$$

A 的特征值 λ_1 与 λ_2 是(与 ω 有关的)复数，它们满足

$$\lambda_1\lambda_2=1, \quad \lambda_1+\lambda_2=2a. \tag{56}$$

此外，任一复数都可写成

$$re^{j\beta}=e^{\alpha}e^{j\beta}=e^{\alpha+j\beta}$$

的形式，其中 α,β 为实数. 我们置

$$\lambda_1=e^{\xi}, \quad \xi=\alpha+j\beta, \tag{57}$$

其中 α,β 为实数. 从(56)式推得

$$\lambda_2=e^{-\xi}, \tag{58}$$

$$a=\frac{1}{2}(e^{\xi}+e^{-\xi})=\cosh\xi. \tag{59}$$

我们不妨设 λ_1 是使 $\alpha\geqslant 0$ 的特征值. 特征值 $\lambda_1=e^{\xi}$ 对应的特征向量空间 $\mathscr{S}_{\lambda_1}$的生成元为

$$\begin{pmatrix} z \\ 1 \end{pmatrix}, \text{ 这里} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix}=e^{\xi}\begin{pmatrix} z \\ 1 \end{pmatrix}.$$

因此

$$az+b=e^{\xi}z,$$

从而

$$z=\frac{b}{e^{\xi}-a}=\frac{b}{\frac{1}{2}(e^{\xi}-e^{-\xi})} \text{ (由(59)式)}$$

$$=\frac{b}{\sinh\xi}, \tag{60}$$

因此，只要 $\sinh\xi\neq 0$ 即 $a\neq\pm 1$ 就可以了. 注意

$$z^2=\frac{b^2}{\cosh^2\xi-1}=\frac{b^2}{a^2-1}=\frac{b^2}{bc}=\frac{b}{c}=\frac{1}{4}Z_1^2+Z_1Z_2, \tag{61}$$

由此，可以取适当的平方根而得到 z. 再注意 $z(=z(\omega))$具有阻抗的量纲. 它称为这个T形网络对于角频率 ω 的特征阻抗.

特征值 $\lambda_2=e^{-\xi}$ 对应的特征空间 $\mathscr{S}_{\lambda_2}$ 由向量 $\begin{pmatrix}-z\\1\end{pmatrix}$ 生成.

我们作矩阵

$$C=\begin{pmatrix}z & -z\\1 & 1\end{pmatrix},\quad 使\quad C^{-1}AC=\begin{pmatrix}e^{\xi} & 0\\0 & e^{-\xi}\end{pmatrix}. \tag{62}$$

所以

$$C^{-1}A^nC=(C^{-1}AC)^n=\begin{pmatrix}e^{n\xi} & 0\\0 & e^{-n\xi}\end{pmatrix},$$

由此不难得到

$$A^n=\begin{pmatrix}\cosh n\xi & z\sinh n\xi\\ \frac{1}{z}\sinh n\xi & \cosh n\xi\end{pmatrix}. \tag{63}$$

现在假定我们取 $z\left(=\frac{b}{\sinh\xi}\right)$为该梯级电路的输出阻抗

$$Z_{出}=z, \tag{64}$$

则

$$E_{n+1}=zI_{n+1},$$

因此,由(54)式

$$\begin{pmatrix}E_1\\I_1\end{pmatrix}=\begin{pmatrix}\cosh n\xi & z\sinh n\xi\\ \frac{1}{z}\sinh n\xi & \cosh n\xi\end{pmatrix}\begin{pmatrix}zI_{n+1}\\I_{n+1}\end{pmatrix}.$$

由其中第二个方程推及

$$\begin{aligned}I_1&=(\sinh n\xi+\cosh n\xi)I_{n+1}\\&=e^{n\xi}I_{n+1}.\end{aligned}$$

把 $\xi=\alpha+j\beta$, $\alpha\geqslant0$ 代入,我们得到

$$I_{n+1}=e^{-n\alpha}(e^{-nj\beta}I_1). \tag{65}$$

这就证明了

(i) 若 $\alpha>0$,电流通过电路时被减弱,并且,当 n 充分大时,

在此梯级电路的第 n 级上，电流可以忽略不计.

(ii) 若 $\alpha=0$，则电流通过此电路后并不减小其振幅.

由于这个原因，α 称为该滤波器电路的衰减常数. β 称为相位常数，因为它给出了每级上相位的变化. ξ 称为传播常数. 这些术语可能会引起误解，因为事实上 α,β 与 ξ 并非真正的常数，而是输入电流角频率 ω 的函数.

条件(i)与(ii)是在(64)的假设下得到的，即假定了输出阻抗等于特征阻抗. 可以证明，在这种情况下，滤波器的工作效率达到极大值.

最后， 我们再把无衰减通过的条件用电路的阻抗表出. 当 $\alpha=0$ 时，方程(59)变成

$$a=\cosh(j\beta)=\cos\beta.$$

再者，当 $\alpha\neq 0$ 时，则 a 不是实数或者是使 $|a|>1$ 的实数（练习 20）. 所以，无衰减通过的条件是 a 为实数且 $-1<a<1$（极限情况 $a=\pm1$ 需要另行讨论）.

因而，只要 a 是实数且 $-1<a<1$，电流就可以无衰减地通过滤波器. 当然，这是对输入电流角频率 ω 的一个限制. 回忆一下

$$a=\frac{\frac{1}{2}Z_1+Z_2}{Z_2}$$

是与 ω 有关的.

例 7　低通滤波器

图 99 表示某个梯级系统的一部分. 求能无衰减地通过该滤波器的电流角频率范围.

对于角频率为 ω 的输入电流，该滤波器的阻抗是

$$Z_1=j\omega L,\qquad Z_2=-\frac{j}{\omega C}.$$

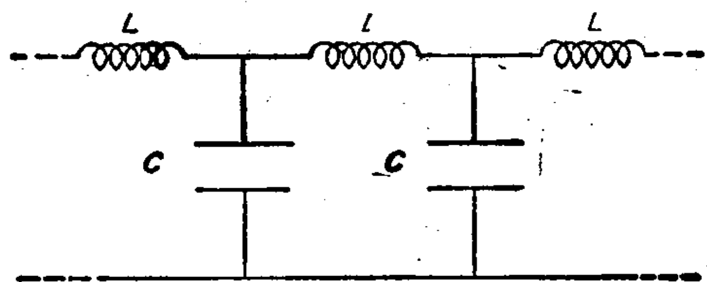


图 99 低通滤波器

从而

$$a=\frac{\frac{1}{2}Z_1+Z_2}{Z_2}=-\frac{1}{2}\omega^2LC+1,$$

这是一个实数.

该滤波器的无衰减通过的频率应满足$|a|<1$,亦即

$$0<\omega^2LC<4, \tag{66}$$

或者说

$$0<\omega<\omega_c, \quad \omega_c=\frac{2}{\sqrt{LC}}.$$

ω_c称为截止频率. 所有频率小于 ω_c 的电流无衰减地通过该滤波器,而所有频率大于 ω_c 的电流被减弱.

特征阻抗 z 由

$$z^2=\frac{1}{4}Z_1^2+Z_1Z_2=-\frac{1}{4}\omega^2L^2+\frac{L}{C}$$

$$=\frac{1}{4}(4-\omega^2LC)L/C$$

给出. 由条件(66)推知,所有能无衰减通过的电流频率都使 z 是实数(一个纯电阻). 因为 z 是随 ω 变化的,所以不可能使所有频率都满足条件(64). 选取效率最高的输出阻抗的方法在一些高级教本中讨论到.例如,可参看 W. R. Smythe, *Static and Dynamic Electricity*, Second Edition, McGraw-Hill, 1950.

用矩阵法讨论电路理论的问题，还可参阅 A. Mary Tropper, *Matrix Theory for Electrical Engineering Students*, Harrap, 1962. 此书中还列出另外一些参考书籍.

练　　习

19. 图 100 表示某个滤波器的一部分，求能无衰减地通过该滤波器的电流角频率分布. 这是一个高通滤波器.

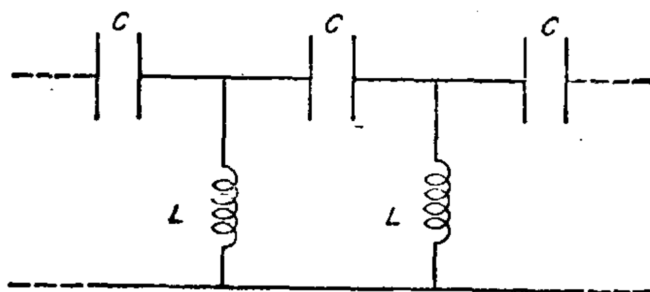


图 100　高通滤波器

20. 设 $a=\cosh(\alpha+j\beta)$，这里 α 与 β 是实数且 $\alpha\neq0$. 证明，如果 a 是实的，则 $|a|>1$.

[提示：用公式

$$\cosh(x+y)=\cosh x\cosh y+\sinh x\sinh y$$

把 $\cosh(\alpha+j\beta)$ 展开.]

附　　录

等价关系

我们常常根据某种性质来对某一数学系统中的元素进行分类，这种方法是很有用的．比如，可以根据三角形的角或面积对三角形进行分类．在第一种分类中，出现在同一类里的二个三角形是相似的（即对应角相等），二个不相似的三角形出现在不同的类里．“相似”这个关系是定义在所有三角形的集合上的等价关系的一个例子．等价关系是相等这个概念的推广，它根据某个性质来衡量相等．

定义 I　设$\mathscr{S}$是一个集合，$\sim$是定义在$\mathscr{S}$的元素上的一个关系，使对$\mathscr{S}$中的任意二个元素a与b，关系$a\sim b$或者成立，或者不成立（写成‘$a\nsim b$’）．当且仅当$\sim$依以下方式把$\mathscr{S}$分成互不相交的类（即没有两个类有公共的元素）时，$\sim$称为等价关系：

(i) 若$a\sim b$，则a与b属于同一类；

(ii) 若$a\nsim b$，则a与b属于不同的类．

例 1　$\mathscr{S}$是所有整数$0, \pm 1, \pm 2, \cdots$的集合．下面是定义在$\mathscr{S}$上的各种关系，哪些是等价关系？对等价关系，叙述所得到的分类情况．

(i) $a\sim b$当且仅当$a=b$．

(ii) $a\sim b$当且仅当$a-b$被2整除．

(iii) $a\sim b$当且仅当$a-b$被3整除．

(iv) $a\sim b$当且仅当$a\leqslant b$．

(v) $a\sim b$当且仅当a与b有大于1的公因子．

(vi) $a\sim b$当且仅当$ab\neq 0$．

(i) ‘相等’是等价关系，对应的分类是平凡分类，即每类中只有一个元素.

(ii) 是等价关系，它把$\mathscr{S}$分成两类，一类是偶数 $0, \pm 2, \pm 4, \cdots$，一类是奇数$\pm 1, \pm 3, \pm 5, \cdots$.

(iii) 是等价关系，它把$\mathscr{S}$分成三类，一类是整数$3k$，一类是整数$3k+1$，一类是整数$3k+2$，所以$41 \sim -4$(因为45被3整除)，41与-4都属于$3k+2$这一类.

(iv) 不是等价关系，否则由定义Ⅰ，我们得到一个分类. 考虑整数5与7，因为$5 \leqslant 7$，5与7在同一类里. 但是$7 \nleqslant 5$，于是7与5不在同一类里，矛盾.

关系～称为对称关系，如果从$a \sim b$可推出$b \sim a$. 关系(iv)因为不是对称关系，所以它不是等价关系.

(v) 不是等价关系. 若不然，我们考虑整数9，15，20. 因为9与15有公因子3，所以它们在同一类里. 同理，15与20在同一类里. 但是9与20没有公因子，所以它们在不同的类里. 矛盾.

关系～称为传递的关系，如果从$a \sim b, b \sim c$可以推出$a \sim c$. 关系(v) 不是传递的，所以它不是等价关系.

(vi) 不是等价关系. 否则我们考虑整数0. 因为$0 \cdot 0 = 0$，得$0 \nsim 0$. 于是由定义Ⅰ，0与0在不同的类里!

定义在集合$\mathscr{S}$上的关系～称为自反的关系，如果对$\mathscr{S}$中所有元素a，有$a \sim a$. 关系(vi)不是自反的，数0与分类不相容，所以它不是等价关系. 注意关系(vi)是对称的，又是传递的.

从上面例子中，可清楚看出等价关系一定是对称的、传递的、自反的. 值得注意的是上面的逆命题也成立.

定理Ⅰ 若关系～是对称的、传递的、自反的，则它是等价关系.

证明 设～定义在集合$\mathscr{S}$上. 我们应该证明这个关系以定义Ⅰ中规定的方式在$\mathscr{S}$上导出一个分类.

用 $\mathscr{C}(a)$代表 $\mathscr{S}$中的使 $x\sim a$ 的元素 x 的集合. 由自反性, a 属于 $\mathscr{C}(a)$,所以集合 $\mathscr{C}(a),\mathscr{C}(b),\cdots$穷尽了 $\mathscr{S}$ 中的所有元素. 我们通过证明

(i) 若 $a\sim b$,则 $\mathscr{C}(a)$与 $\mathscr{C}(b)$是同一个集合;

(ii) 若 $a\nsim b$,则 $\mathscr{C}(a)$与 $\mathscr{C}(b)$不相交.

来证明本定理.

(i) 设 $a\sim b$,考虑 $\mathscr{C}(a)$中的 x,即 $x\sim a$. 由传递性,从 $a\sim b$ 推出 $x\sim b$,所以 x 属于 $\mathscr{C}(b)$.

由对称性,$a\sim b$ 意味着 $b\sim a$,如上同理可证属于 $\mathscr{C}(b)$ 的每个 x 属于 $\mathscr{C}(a)$,所以 $\mathscr{C}(b)$与 $\mathscr{C}(a)$是同一个集合.

(ii) 设 $a\nsim b$, 而 x 同时属于 $\mathscr{C}(a)$ 与 $\mathscr{C}(b)$. 于是 $x\sim a$, $x\sim b$,由对称性,$a\sim x$, 所以从传递性可推出 $a\sim b$, 与假设矛盾. 定理证毕.

本书讨论二个重要的等价关系,一个是行等价 $\underset{r}{\sim}$,这是定义在所有 $m\times n$ 矩阵的集合上的关系(2.3节); 一个是相似, 这是定义在所有 $n\times n$ 矩阵的集合上的关系(13.3节).

关系 $\underset{r}{\sim}$ 在 $m\times n$ 的矩阵集合上导出一个分类, 使当且仅当二个矩阵所对应的线性方程组有相同的通解, 这二个矩阵属于同一类. 每类包含唯一的简化阶梯形矩阵,这个矩阵可以方便地用来代表这个类.

相似关系是根据 $n\times n$ 矩阵变换是否在坐标系的变换下相关来对 $n\times n$ 矩阵变换进行分类. 如有可能,取对角矩阵作为一个类的代表是很方便的. 但是,某些类不包含任何对角矩阵. 例如, 不存在与切变矩阵

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

相似的对角矩阵.

练　　习

下面定义在集合 $\mathscr{S}$ 上的关系中，哪些关系是(a)对称的，(b)传递的，(c)自反的？对每个等价关系，试描述由这个等价关系在集合 $\mathscr{S}$ 上所定义的分类的情况。

1. $\mathscr{S}$ 是整数集合，关系是：

(i) $\neq$，不等于；

(ii) $<$，小于；

(iii) 整除。

2. $\mathscr{S}$ 是平面上直线的集合，关系：平行。

3. $\mathscr{S}$ 是变量 x 的实系数多项式，当且仅当 $p_1(x)-p_2(x)$ 被 x^2+1 整除时，$p_1\sim p_2$。

4. $\mathscr{S}$ 是一个群，它有子群 $\mathscr{H}$，关系 $\sim$ 是：当且仅当 ab^{-1} 属于 $\mathscr{H}$ 时 $a\sim b$。

5. $\mathscr{S}$ 是 $n\times n$ 实矩阵集合。关系是：

(i) $\sim$，当且仅当存在矩阵 P 与 Q，使 $B=PAQ$，称 $A\sim B$。

(ii) $\sim$，当且仅当存在非奇异矩阵 P 与 Q，使 $B=PAQ$，称 $A\sim B$。

6. $\mathscr{S}$ 是所有群的集合，关系 $\cong$ 是：当且仅当 G_1 与 G_2 是同构时，$G_1\cong G_2$。

习题答案与注释

第一章

1. 封闭律不成立。例如，y 与 $2y$ 不可能都是这个微分方程的解。注意法则 IV 也不成立：零函数 $y(t)=0$（对一切 t）也不是这个微分方程的解。

10. 设 $\mathbf{s}$ 属于 $\mathscr{S}$，则 $(-1)\mathbf{s}=-\mathbf{s}$ 及 $0\,\mathbf{s}=\mathbf{0}$ 也属于 $\mathscr{S}$。于是，$\mathscr{S}$ 中的向量满足法则 IV 与 V。因为法则 II，III 与 VII 在 $\mathscr{V}$ 中成立，所以它们在 $\mathscr{S}$ 中也成立。

11. 不是；二个长度是单位的向量之和不一定是长度为单位的向量。

14. 图 101 的三角形 ABC 说明此人以每小时 12 英里骑自行车时的情况．$\overrightarrow{AC}=\mathbf{w}$ 是真正的风速，$\overrightarrow{BC}=\mathbf{v}$ 代表向正南每小时 12 英里，$\overrightarrow{AB}=\mathbf{w}-\mathbf{v}$．与此类似，三角形 $A'BC'$ 说明了第二种情况。因为 AC 与 $A'C'$ 互相平行且

相等，所以 $ACC'A'$ 是平行四边形．于是 AA' 是铅垂方向的，长度是 9 个单位，由于角 ABA' 等于 $45°$，故三角形 $A'AB$ 是等腰三角形，$AA'=AB=9$．所以

$$\|\mathbf{w}\|^2=9^2+12^2=15^2.$$

风从北偏西 α 角方向以每小时 15 英里速度吹来，这里 $\tan\alpha=\frac{3}{4}$．

15．$3\overrightarrow{AD}$．

17．$\mathbf{a}+\mathbf{c}-\mathbf{b}$．

18．用 $\mathbf{a}$，$\mathbf{b}$，$\mathbf{c}$，$\mathbf{d}$ 记这个四面体的四个顶点的位置向量交点是 $\frac{1}{4}(\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d})$．

19．把原点取在这个正方形中心 O 处．关于此原点，质量中心的位置向量是 $\frac{1}{10}(\mathbf{a}+2\mathbf{b}+3\mathbf{c}+4\mathbf{d})=\frac{1}{5}(\mathbf{c}+\mathbf{d})$．质量中心在过 O 点的 CD 的垂线上，离 O 点 $\frac{1}{5}$ 个单位．

20．$2\mathbf{v}=(4,-2)$，$3\mathbf{w}=(3,-3)$，$2\mathbf{v}-3\mathbf{w}=(1,1)$．

21．$\left(-1,\ \frac{1}{3}\right)$．

22．$\left(\frac{1}{2},\ \frac{7}{10}\right)$．

23．$(1,1,0)$．

27．零向量 $(0,0,0,0)$ 不是解向量．

28．解向量全体成 $\mathscr{C}^2$ 的一个子空间，它由所有形如 $c(-1,1+i)$ 的向量组成，c 是复数纯量．

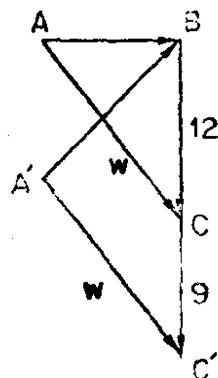


图 101　骑车问题

29．(i) 子空间；(ii) 不是子空间；(iii) 子空间．

第二章

2．$x_1=-\frac{1}{4}$，$x_2=0$，$x_3=0$ 是方程组(7)的解，但不是(6)的解．

4．设 $x_i=a_i(i=1,\cdots,n)$ 是这个齐次方程组的非平凡解．则 $x_i=ka_i$ $(i=1,\cdots,n)$ 也是这个方程组的解．

5．$I_1=4$ 安培，$I_2=3$ 安培，$I_3=1$ 安培．

6．它的简化阶梯形矩阵是 $\begin{pmatrix}1 & 0 & 7 & 8\\ 0 & 1 & -5 & -8\end{pmatrix}$，对应的线性方程组

的通解是 $x_3=k$(任意)，$x_1=8-7k$，$x_2=-8+5k$.

11．可用一个类型II的变换，用1乘某行，便可把矩阵 A 化成本身．

12．用类型II和类型III的变换可把这个矩阵的第一行变成 $(k \quad 2k \quad l \quad k+2l)$ 的形式，但不能变到任何其它形式．

13．设这两个简化阶梯形矩阵是 A 与 B．容易证明如果 A 与 B 的对应行的首元素出现在不同位置处，则 A 与 B 不是行等价的．

现设

$$A=\begin{pmatrix}1 & 0 & a_{13} & a_{14}\\0 & 1 & a_{23} & a_{24}\end{pmatrix},\quad B=\begin{pmatrix}1 & 0 & b_{13} & b_{14}\\0 & 1 & b_{23} & b_{24}\end{pmatrix}.$$

这里还可设 $a_{13}\neq b_{13}$．用类型II与类型III的变换可把 A 的第一行变成 $(k \quad l \quad ka_{13}+la_{23} \quad ka_{14}+la_{24})$ 的形式，但不能变成其它形式．但是 B 的任何行都不是这种形式，所以 A 不行等价于 B．对 A 与 B 可能有的其它形式，用同样方法可证．

15．(i) $x_3=k$(任意)，$x_4=l$(任意)，$x_1=\dfrac{1}{3}(13+3k-5l)$，

$x_2=\dfrac{1}{3}(7-2l)$；

(ii) $x_1=\dfrac{1}{2}(v_1+v_2-v_3)$，$x_2=\dfrac{1}{2}(v_3+v_1-v_2)$，$x_3=\dfrac{1}{2}(v_2+v_3-v_1)$；

(iii) $x_1=3$，$x_3=k$(任意)，$x_5=l$(任意)，$x_2=1-2k-3l$，$x_4=-1-l$．

16．$x_2=c$(任意)，$x_1=1+i-ci$，$x_3=2$．

17．$x_1=1$，$x_2=0$，$x_3=4$，$x_4=0$．由前面三个方程组成的方程组的通解是 $x_4=k$(任意)，$x_1=1+\dfrac{1}{4}k$，$x_2=-k$，$x_3=4+\dfrac{9}{4}k$．

19．例如，方程组 $x_1+x_2=3$，$x_1-x_2=1$，$x_1+2x_2=4$ 有唯一解 $x_1=2$，$x_2=1$；方程组 $x_1+x_2=3$，$2x_1+2x_2=6$，$3x_1+3x_2=9$ 有无数多个解 $x_2=k$(任意)，$x_1=3-k$．

23．方程组 $x+2y=0$，$x+(2-\varepsilon)y=-1$ 有通解 $x=-2/\varepsilon$，$y=1/\varepsilon$．对 ε 的小变化，这个通解是很敏感的．

24．方程组 $x-2y=0$，$x+(2-\varepsilon)y=-1$ 有通解 $x=-2/(4-\varepsilon)$，$y=-1/(4-\varepsilon)$．

25．通解是

$$x_1=\frac{a_{12}}{a_{11}a_{22}-a_{12}a_{21}},\quad x_2=\frac{-a_{11}}{a_{11}a_{22}-a_{12}a_{21}}.$$

27．证明$(a_{11}a_{22}-a_{12}a_{21})x_2=-a_{11}$，$(a_{11}a_{22}-a_{12}a_{21})x_1=a_{12}$．所以如果$a_{11}\neq 0$或$a_{12}\neq 0$时，这个方程组是不相容的．若$a_{11}=a_{12}=0$以及$a_{21}$与$a_{22}$中至少有一个不是零时，这个方程组有无数个解．最后，如果对所有的i, j都有$a_{ij}=0$，则此方程组显然是不相容的．

第三章

1．(i) $2\mathbf{v}+3\mathbf{w}$；(ii) $\mathbf{w}-2\mathbf{v}$；(iii) $-3\mathbf{w}$．

所有线性组合都是唯一的．

2．(i) 不能生成；(1,0,0) 不是这些向量的线性组合；(ii) 不能生成，(iii) 能生成；(iv) 能生成．

3．$\mathbf{a}=1\,\mathbf{b}+1\,\mathbf{c}+0\,\mathbf{d}$．

4．向量(1,0,0,0)不能表示成 **a**、**b**、**c**、**d** 这四个向量的线性组合．

5．例如，(1,0,0,0)．

7．(i) 能生成；(ii) 不能生成．

8．(i) $2\mathbf{v}_1-\mathbf{v}_2=0$；

(ii) $\mathbf{v}_1-\mathbf{v}_2+\mathbf{v}_3=0$；

(iii) 线性无关；

(iv) $\mathbf{v}_1-\mathbf{v}_2-\mathbf{v}_3=0$．

9．以所给四个向量为行的 4×3 矩阵A等价于一个阶梯形矩阵，这个梯矩阵的最后一行不全是零．所以A的各行是线性无关的．

10．(i) 线性无关；(ii) $2\mathbf{v}_3=\mathbf{v}_2-i\mathbf{v}_1$．

12．逆命题正确．

14．这r个向量中的一个向量可以写成其余$r-1$个向量的线性组合．这$r-1$个向量生成同一个向量空间．

15．(i) 不是一组基；(ii) 不是一组基；(iii) 基；$(2,3,0)=\frac{5}{2}\mathbf{v}_1-\frac{3}{2}\mathbf{v}_2+\mathbf{v}_3$；(iv) 不是一组基．

16．(i)，(ii) 及 (iv) 不是基；

(iii) 基；$(2+x)=\frac{5}{2}(1+x+x^2)-\frac{3}{2}(1-x+x^2)+(1-x-x^2)$．

17．3．

18. 这个集合不包含零. (也请注意,两个 3 次多项式的和可能是小于 3 次的多项式.)

19. 例如,(1,0,0,0).

21. 设某个向量可以用两种方法写成 $\mathbf{v}_1, \cdots, \mathbf{v}_n$ 的线性组合,则可以得到这 n 个向量的一个非平凡的线性关系.它的逆命题是正确的,因为如果零向量有唯一的表达式 $\mathbf{0}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 则可推出 $c_1=\cdots=c_n=0$, 所以 $\mathbf{v}_1, \cdots, \mathbf{v}_n$ 没有非平凡的线性关系.

22. 例如,$(2,-1,3)=0\,\mathbf{v}_1+2\,\mathbf{v}_2-\mathbf{v}_3+3\,\mathbf{v}_4=2\,\mathbf{v}_1-3\,\mathbf{v}_3+3\,\mathbf{v}_4$.

23. (i) 基;

(ii) 不是基: 多项式 $p(x)=1$ 不能写成有限个题中所写出的多项式的线性组合.

(iii) 基.

24. 由公式 $2\sin qx\sin rx=\cos(q-r)x-\cos(q+r)x$ 得, 当 $q\neq r$ 时, $\int_{-\pi}^{\pi}\sin qx\sin rx\,dx=0$.

第四章

1. $\|\mathbf{v}\|=\sqrt{6}$. 它的方向余弦: $\cos\alpha=1/\sqrt{6}$, $\cos\beta=-1/\sqrt{6}$, $\cos\gamma=2/\sqrt{6}$. 正规化 $\mathbf{v}$ 可得向量

$$\frac{1}{\sqrt{6}}\mathbf{v}=\left(\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}\right).$$

3. $\overrightarrow{PQ}=(1,-1,2)=\overrightarrow{SR}$, $\overrightarrow{PS}=(-2,1,-1)=\overrightarrow{QR}$. 对角线的长是 $PR=\sqrt{2}$, $QS=\sqrt{22}$.

4. (i) $\cos\theta=-\frac{4}{9}$; (ii) $\cos\phi=\frac{4}{9}$, $\theta+\phi=\pi$.

5. $PQ=PR=3$, $\cos P\hat{Q}R=-\frac{4}{9}$, QR 的中点是 $M\left(1,\frac{1}{2},\frac{1}{2}\right)$.

6. 与 $\mathbf{v}$ 与 $\mathbf{w}$ 正交的单位向量是: $\pm\frac{1}{3}(2,-2,1)$.

9. 与 $\mathbf{b}$ 正交的,形式是 $\mathbf{a}-k\mathbf{b}$ 的最短向量由

$$k=\frac{(\mathbf{a}\cdot\mathbf{b})}{\|\mathbf{b}\|^2}$$

给出.

10. 对角线互相垂直.（这个平行四边形是菱形.）

11. 当且仅当$\|\mathbf{a}\|=\|\mathbf{b}\|$时，$(\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}+\mathbf{b})=\|\mathbf{a}\|^2-\|\mathbf{b}\|^2=0$。当且仅当平行四边形是菱形时，平行四边形的对角线互相垂直.

12. 考虑$(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b})$。这是毕达哥拉斯定理及其逆定理.

13. 考虑$(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b})+(\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})$。

14. $\mathbf{F}\cdot\mathbf{s}$。

16. (i) $\cos\theta=-\frac{33}{65}$； (ii) 单位向量$\pm\frac{1}{5}(4,-3)$。

17. 第一个平面：单位法向量$\mathbf{u}_1=\frac{1}{3}(1,2,-2)$，原点到法向量的距离是 $p_1=\frac{1}{3}$。

第二个平面：$\mathbf{u}_2=\frac{1}{3}(2,1,2)$，$p_2=1$。因为$\mathbf{u}_1\cdot\mathbf{u}_2=0$，所以它们互相垂直.

18. $\mathbf{u}_1=\frac{1}{3}(1,2,-2)$，$p_1=1$。第二个方程是$-x-2y+2z=3$，所以$\mathbf{u}_2=-\mathbf{u}_1$，$p_2=1$。平面之间的垂直距离是2个单位.

19. $x+2y+2z=1+2.2+2(-2)=1$。

20. $\frac{2}{3}$.

21. $2x-2y-3z=14$，$p=\frac{14}{\sqrt{17}}$，$\mathbf{u}=\frac{1}{\sqrt{17}}(2,-2,-3)$。

24. (i) 二维波以速率$v/5$沿$\mathbf{u}=\left(\frac{3}{5},-\frac{4}{5}\right)$方向传播. 波长$2\pi/5$；

(ii) $U=-a\sin(-3x+4y-vt)$表示速率为$v/5$，方向是$-u=\left(-\frac{3}{5},\frac{4}{5}\right)$的波.

25. $\frac{1}{\sqrt{3}}(-1,1,1)$。

26. $-x+y+z=1$。

27. $7/2\sqrt{3}$.

28. 不能；例如$\mathbf{i}\times\mathbf{j}=(\mathbf{i}+\mathbf{j})\times\mathbf{j}$。

33．$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}$。

34．当且仅当 **a**、**b**、**c** 三个向量共面时，$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})=0$。

35．(i) $(2,-4,-3)$；(ii) $(-1,-3,-1)$；

(iii) $(0,0,0)$。B 点在 **F** 的作用线上。

36．$\mathbf{w}=(2,4,4)$。A 的速度是$(0,12,-12)$等等，因为 D 在轴上，D 的速度是零。

37．$\mathbf{w}=c(1,1,1)$，c 是常数。P 的瞬时速度是$(c,-c,0)$。

第五章

2．(i) 不对称；

(ii) 对称，但是当 $\mathbf{v}=(0,1,1)$时，$\langle\mathbf{v},\mathbf{v}\rangle$是负的；

(iii) 除了 $\mathbf{v}=(0,1,1)$外，$\langle\mathbf{v},\mathbf{v}\rangle$取非负值；

(iv) 内积；

(v) 不是双线性的。

3．(i)、(iii)、(iv)及(vi)定义了内积。

4．所有向量的长度是 2。

11．设 f 与 g 是偶函数，$h(x)=f(x)g(x)$。则

$$h(-x)=f(-x)g(-x)=f(x)g(x)=h(x).$$

12．$\sin kx$ 是奇函数，$\cos kx$ 是偶函数。

15．(i) 令 $f(x)=\sqrt{x}$，$g(x)=\sqrt{\sin x}$；

(ii) 令 $f(x)=1$，$g(x)=\sqrt{\sin x}$。

16．0.91。

18．当且仅当 $f=\theta$(零向量)，或 $g=kf$，这里 k 是非负纯量这几种情况，三角形不等式变成等式。

21．不成立；例如可以考虑 $g=-f\neq\theta$ 的情况。此时，$\|f-g\|=2\|f\|>0$，但是$\|f\|-\|g\|=\|f\|-\|f\|=0$。

22．$(16,-11,-1)$。

23．$\mathbf{u}_3=\pm\left(-\sqrt{\frac{1}{2}},\ \frac{1}{2},\ \frac{1}{2}\right)$。

24．$\pm\frac{1}{\sqrt{38}}(3,-5,2)$。

25．对 $i\neq k$，用等式$\langle f_i,f_k\rangle=0$；对 $i=k$，用等式$\langle f_k\cdot f_k\rangle=1$。

26. 分量是$\frac{1}{3}$，$\frac{5}{3}$，$-\frac{1}{3}$.

27. 由 5.4 节例 13，这些多项式是线性无关的。

$$x^3+x^2+1=\frac{2}{3}P_3(x)+\frac{2}{3}P_2(x)+\frac{3}{5}P_1(x)+\frac{4}{3}P_0(x).$$

28. 设f是次数≤3 的多项式，则$f=\sum_{i=0}^{3}c_iP_i$。若对所有的i，$\langle f,P_i\rangle=0$，则对所有i，$c_i=0$，所以$f=\theta$. 为了证明本题的最后部分，只需注意每个次数≤2 的多项式可以写成P_0,P_1,P_2的线性组合。

29. $L_3(x)=6-18x+9x^2-x^3$；$p=9L_0-22L_1+10L_2-L_3$。

30. 具有给定形式的向量中的最短长度向量垂直于$\mathbf{u}_1$及$\mathbf{u}_2$。所以，$c_i=\langle \mathbf{v},\mathbf{u}_i\rangle(i=1,2)$。

37. (i) 对$n\geqslant 1$，$\langle P_n,P_0\rangle=0$。

(ii) 对$n\geqslant 1$，$\int_{-\infty}^{\infty}e^{-x^2}H_n(x)dx=0$，等等.

39. 例如，$\frac{1}{2}(1,1,-1,-1)$ 与 $\frac{1}{2}(1,-1,-1,1)$。

42. 在酉空间里$\langle f,g\rangle+\langle g,f\rangle=\langle f,g\rangle+\overline{\langle f,g\rangle}=2(\langle f,g\rangle$的实部)。甚至$f$与$g$不是正交时，它也可能是零.

45. 考虑练习 44 中的公式，$k\|g\|^2=\overline{\langle f,g\rangle}$。

46. 这是许瓦尔兹不等式的特例。

第六章

1. 对称变换是：

(i) 关于正五边形中心的旋转，旋转角是$2\pi/5$的倍数，以及关于过这个中心及一个顶点的直线的反射；

(ii) 恒等变换，关于正六边中心的π角的旋转，关于两个相对顶点连线的反射；

(iv) 恒等变换与关于这个平行四边形中心的π角的旋转；

(v) 恒等变换，关于椭圆中心的π角的旋转，关于主轴的反射；

(vi) 恒等变换与关于抛物线的轴的反射。

6. 当且仅当平移T与S的反射直线L平行时，S与T可以交换。

7. 关于不同中心的旋转不可交换。（考虑其中一个旋转中心在乘积下的象.）

12．(i) 由$-\mathbf{a}$定义的一个平移；(ii) 同一个反射。

13．例如，反射以及转π角的旋转。

19．$R^{74}=R^2$，$R^{-39}=R^3$。

20．(i) 5；(ii) 2。

21．若$T^n=1$，则$T^{-1}=T^{n-1}$。逆命题不成立，平移有逆，但它不是周期的。

22．$(ST)^2=STST=S(ST)T=S^2T^2$等等．(可对$n$进行归纳来证明一般情况．) 此公式对不可交换的变换不成立。例如，可考虑关于两条平行线的反射S_1及S_2。此时，$S_1^2S_2^2=1$，但是$(S_1S_2)^2\neq1$。

23．(i) 关于由L平移$-\frac{1}{2}\mathbf{a}$所得直线的反射；及(ii) 一个滑移反射。

24．关于过O点的，与L倾成$-\frac{1}{2}\alpha$角的直线的反射；及(ii) 一个滑移反射。

26．(i) 二倍于镜面间距离的平移；

(ii) 旋转角是镜面夹角的二倍的旋转。

27．$(S_1S_2)^{-1}=S_2^{-1}S_1^{-1}=S_2S_1$。

29．$T^{-1}(ST)T^{-1}=T^{-1}(TS)T^{-1}$，等等。

31．除了练习 30 中讨论的情况外，乘积是旋转。为了说明情况(ii)，把其中的一个旋转写成一个平移与一个关于另一个旋转中心的旋转之乘积．当且仅当二个旋转($\neq1$)的中心重合时，它们可以交换。

33．关于垂直于滑移反射轴的直线的反射。

34．中心在滑移反射轴上的一个半周旋转。

36．旋转轴是二个反射平面的交线，旋转角是平面夹角的二倍。

39．乘积是使一个点固定的正等距变换。

40．一只左手手套。

41．一个反射。

42．(i) 关于过此正方体中心的一条水平轴的半周旋转；

(ii) 关于过此正方体中心的水平平面的反射；

(iii) 没有非平凡的对称。

43．取图 50 中的正方体，用红色盖掉蓝色，白色盖掉绿色。把正方形

1、2、3、4 分别涂上白、红、白、红四种颜色。

45. (i) n; (ii) $2n$.

第七章

1. (ii)与(iv)是循环群; (i)与(iii)不是群。

2. 由定义关系式 (2), $(SC_n)^2=SC_nSC_n=S^2C_n^{-1}C_n=1$. 所以$(SC_n)^{-1}=SC_n$. 由同一个关系式, $(C_n^r)^{-1}=C_n^{n-r}$, 故$(SC_n^r)^{-1}=SC_n^r$.

3. (i) S; (ii) S; (iii) SC_4.

4. 若 $ag=bg$, 则$(ag)g^{-1}=(bg)g^{-1}$.

5. 不成立; 令 $a=b=C_4$, $c=d=C_4^2$, $g=S$.

6. (i), (ii)与(iv).

8. 若 g 的周期是 n, 则 g^r 的周期是 n/d, 这里 d 是 n 与 r 的最大公约数。

10. 三个 4 阶子群, 一个由 C_4 生成, 一个由 C_4^2 与 S 生成, 一个由 C_4 与SC_4 生成; 五个 2 阶子群, 分别由 C_4^2, S, SC_4, SC_4^2, SC_4^3 生成。

11. 此命题对交换群成立, 但一般情况不成立。例如, 在 $\mathscr{D}_4$ 中, $S^2=(SC_4)^2=1$, 但$(S\cdot SC_4)^2\neq1$.

12. $\mathscr{D}_4$ 的子群由 C_4^2 与 S 生成。

13. 设在某个同构下, $g\leftrightarrow g'$, $h\leftrightarrow h'$, 则 $gh\leftrightarrow g'h'$, $hg\leftrightarrow h'g'$. 所以, 当且仅当 $g'h'=h'g'$时, $gh=hg$.

14. 同构是一一对应。

15. 用练习 13.

16. (i) $g=ge\leftrightarrow g'e'=g'$;

(ii) $e=g^{-1}g\leftrightarrow e'=(g^{-1})'g'$. 所以$(g^{-1})'=(g')^{-1}$.

17. 子群(ii)与(iii)是同构的。

18. (i) (1352)(47); (ii) (124)(567).

19. (i) (14762); (ii) (12345); (iii) (ab).

23. $\pi^2=(135)(246)$, $\pi^3=(14)(25)(36)$. n-循环的同期是 n.

24. m 与 n 的最小公倍数。

25. $\pi\pi'$的周期是 3. 请注意 π 与 π' 有一个公共符号。

26. 每个 4 次置换可写成

$$\pi=\begin{pmatrix}1 & 2 & 3 & 4\\ a_1 & a_2 & a_3 & a_4\end{pmatrix},$$

这里 a_1, a_2, a_3, a_4 是 1，2，3，4 的某个排列，考虑不同的排列的个数.

27．不是循环群.

28．(ii)把正六边形的六个顶点依次编成 1，2，…，6．则用 (123456) 表示 C_6，(26)(35)表示关于过顶点 1 与顶点 4 的直线的反射 S.

29．C_4^2 表示成(1562)(3487).

30．24.

32．旋转：(142)，(13)(24)；反射：(13)，(24)；伪旋转：(1324)，(1234).

33．对称变换：(i)，(iii)，(iv)与(v).

34．(i) 由(13)生成；(ii) 由(13)与(1234)生成.

35．例如：$x_1x_2x_3$ 以及 $x_1+x_2+x_3$.

36．此多项式的对称变换与顶点被适当编号的正方形的对称变换可用相同的置换来表示.

37．(i) 奇置换；(ii) 偶置换.

38．用 7.5 节定理 I 及公式$(ab)=(1\,a)(1\,b)(1\,a)$.

39．奇置换的集合不是群.

41．$\mathscr{G}=gp\{c^3\}\times gp\{c^5\}$.

42．不能；这样的直积中的每个元素的阶都≤ 6.

43．$\mathscr{T}$ 的 12 个元素中只有六个能写成 hk 的形式.

第八章

5．G, G' 是二个滑移反射，它们轴的夹角是 α．当 $\alpha=0$ 时，它们的乘积是平移，否则是旋转．证明：写

$$GG'=(TS)(S'T')=T(SS')T',$$

用 8.1 节练习 4 及 6.3 节例 11.

6．(i) 反射与 (ii) 滑移反射，它的轴垂直于已知的滑移反射的轴.

9．不是；见练习 4.

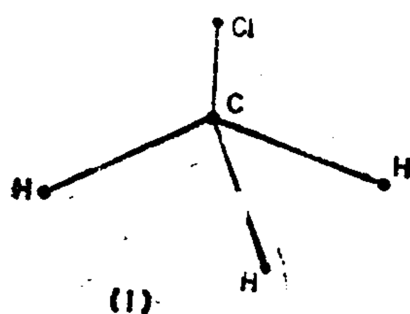
20．(i) $\mathscr{C}_2$；(ii) $\mathscr{C}_{1h}(=\mathscr{C}_s)$；(iii) 只有恒等变换.

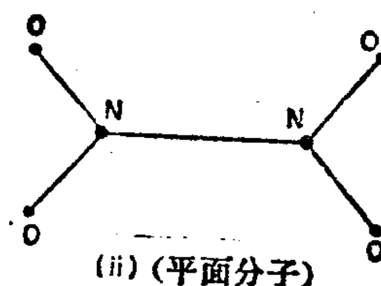
22．(i) 存在把其中一个顶点变到八个可能位置之一的旋转．而且，每个顶点上通过一根 3 次轴，因此有 $8\times 3=24$ 个旋转对称；

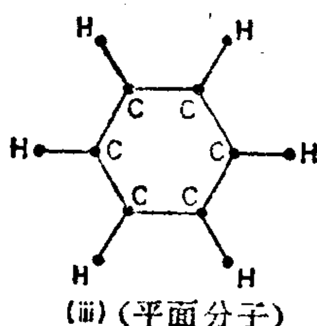
(ii) 完全对称群由旋转与一个水平反射 σ_h 生成.

23．$\mathscr{C}_{1h}$.

24．见图 102.







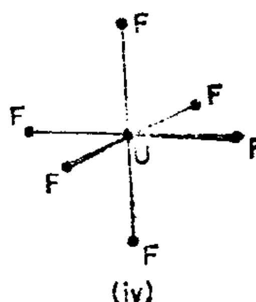


图 102 第八章练习 24 的解答

第九章

1．注意到 $0\mathbf{v}=0$，对 $k=l=0$ 用公式(9.3)，等等。

2．考虑 $R(-\mathbf{v}+\mathbf{v})$。

5．依次用 S 是线性变换与 R 是线性变换这两个事实，证明

$$(RS)(k\mathbf{v}+l\mathbf{w})=k(RS)\mathbf{v}+l(RS)\mathbf{w}.$$

6．$T0\neq 0$。再用练习 1。

10．(i) 关于直线 $x_2=x_1$ 的反射；

(ii) 膨胀；

(iii) 关于 O 的半周旋转；

(iv) 关于 O 点逆时针旋转 α 角，$\tan\alpha=\frac{4}{3}$；

(v) 关于 O 点顺时针旋转 α 角；

(vi) 切变；

(vii) 到 $O2$-轴的正交射影。

13．(i) 关于 Oz 轴的旋转；

(ii) 关于 Oy 轴的旋转；

(iii) 关于 xy 平面的反射。

15. (i) $AB=\begin{pmatrix}-10 & -20\\ 5 & 10\end{pmatrix}$, $BA=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$;

(ii) $AB=BA=\begin{pmatrix}\cos\alpha\cos\beta-\sin\alpha\sin\beta & -(\sin\alpha\cos\beta+\cos\alpha\sin\beta)\\ \sin\alpha\cos\beta+\cos\alpha\sin\beta & \cos\alpha\cos\beta-\sin\alpha\sin\beta\end{pmatrix}$.

16. AB表示一个$(\alpha+\beta)$角的旋转．因此

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta,\text{ 等等。}$$

18. 重复数次正交射影等于一个正交射影．

19. $I=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$.

20. 逆时针旋转 $2(\beta-\alpha)$角．

21. $AB=\begin{pmatrix}0 & -3 & -2\\ 0 & -6 & -4\\ 0 & 9 & 6\end{pmatrix}$, $BA=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$.

22. $A\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}=\begin{pmatrix}a_{11}\\ a_{21}\\ a_{31}\end{pmatrix}$,等等， $A\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}a_{11}x_1+a_{12}x_2+a_{13}x_3\\ a_{21}x_1+a_{22}x_2+a_{23}x_3\\ a_{31}x_1+a_{32}x_2+a_{33}x_3\end{pmatrix}$.

25. $P'(\sqrt{3}, 3)$.

26. 对 $\mathbf{x}=\begin{pmatrix}k\\ k\sqrt{3}\end{pmatrix}$，一切$k$，$A\mathbf{x}=\mathbf{x}$. 点$P(k,k\sqrt{3})$在射影的直线上．对 $\mathbf{x}=\begin{pmatrix}l\sqrt{3}\\ -l\end{pmatrix}$，一切$l$，$A\mathbf{x}=0$. 点$Q(l\sqrt{3}, -l)$在过点$O$所引的射影直线的垂线上．

30. $x_1=4z_1+6z_2$，$x_2=2z_1$。

第十章

1. $X=\begin{pmatrix}1-d & b\\ \dfrac{d(1-d)}{b} & d\end{pmatrix}$，$b\neq0$，及

$X=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$, $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, $\begin{pmatrix}1 & 0\\ c & 0\end{pmatrix}$ 或 $\begin{pmatrix}0 & 0\\ c & 1\end{pmatrix}$，$c$任意．

2. $X=I, -I$, $\begin{pmatrix}1 & 0\\ c & -1\end{pmatrix}$, $\begin{pmatrix}-1 & 0\\ c & 1\end{pmatrix}$，$c$任意，及

$$X=\begin{pmatrix} k & b \\ \frac{1-k^2}{b} & -k \end{pmatrix},\ b\neq 0,\ k\ 任意.$$

4．不能；例如在练习 3 中，$AB=AI$ 并不意味着 $B=I$。

5．不能；例如当 $B=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$ 及 $X=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$ 时，$BB=BI$，但是 $BXB\neq BXI$。

6．$3A=\begin{pmatrix}6 & 9 & 3\\ 0 & -3 & 3\end{pmatrix}$，$3A+2B=\begin{pmatrix}8 & 7 & 3\\ 4 & -1 & 7\end{pmatrix}$，等等。

7．$(5A)\mathbf{x}=5\mathbf{x}'$。

8．这个矩阵是乘积

$$\begin{pmatrix}5 & 0\\ 0 & 5\end{pmatrix}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix},$$

这里 $\cos\alpha=\frac{3}{5}$。

10．一般不成立．正确的表达式是 $(A+B)^2=A^2+AB+BA+B^2$。

11．$2x_1+x_2+2x_3=0$。$(A\mathbf{x}=\mathbf{x}.)$

13．$(A-B)^3=A^3-A^2B-ABA-BA^2+AB^2+BAB+B^2A-B^3$。

14．$A\mathbf{x}=\begin{pmatrix}-3\\ 2\\ 0\end{pmatrix}$，　$(A\mathbf{x})'=(-3\quad 2\quad 0)$。

17．$A'A=\begin{pmatrix}5 & 0 & 5\\ 0 & 5 & 5\\ 5 & 5 & 10\end{pmatrix}$。

18．用公式(13)，$(A'A)'=A'(A')'=A'A$。

19．A'A的 jj-元素是$(a_{1j}^2+a_{2j}^2+\cdots+a_{mj}^2)$。

21．(i) $\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$；　(ii) $\begin{pmatrix}0 & 0 & \frac{1}{2}\\ 0 & 4 & 0\\ \frac{1}{2} & 0 & 0\end{pmatrix}$；

(iii) $\begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & \frac{1}{2} \\ -2 & \frac{1}{2} & 2 \end{pmatrix}$.

22. (i) $x_1y_1+x_2y_2-x_1y_2-x_2y_1$;

(ii) $4x_2y_2+\frac{1}{2}(x_1y_3+x_3y_1)$;

(iii) $x_1y_1+2x_3y_3-2(x_1y_3+x_3y_1)+\frac{1}{2}(x_2y_3+x_3y_2)$.

23. $\langle \mathbf{x}, A\mathbf{y}\rangle = x_1y_1+2x_2y_2+2x_1y_2-x_2y_1=\langle A'\mathbf{x},\mathbf{y}\rangle$,

$\langle A\mathbf{x},\mathbf{y}\rangle = x_1y_1+2x_2y_2-x_1y_2+2x_2y_1$.

24. 除(v)外全是.

25. 矩阵(i)是斜称的.

26. $(A+A')'=A'+A$, $(A-A')'=A'-A=-(A-A')$.

27. $A=\frac{1}{2}(A+A')+\frac{1}{2}(A-A')$.

28. $\mathbf{x}'C\mathbf{x}$ 是一个数（1×1 矩阵），等于它自己的转置．但 $(\mathbf{x}'C\mathbf{x})'=\mathbf{x}'C'\mathbf{x}=-\mathbf{x}'C\mathbf{x}$. 可得所需结论.

29. (i)与 (iii).

30. (i)与 (ii) $6+3i$; (iii)与(v) $8+3i$; (iv) $6+3i$.

$\langle \mathbf{x}, A\mathbf{x}\rangle=9$, $\langle \mathbf{x}, B\mathbf{x}\rangle=6+i$.

第十一章

1. (i) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; (iv) $\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$;

(v) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; (vi) $\begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$.

2. $A^{-1}(AB)=A^{-1}(AC)$, 所以 $B=C$.

3. $(AB)^{-1}=B^{-1}A^{-1}$.

4. $(A_1A_2\cdots A_k)^{-1}=A_k^{-1}\cdots A_2^{-1}A_1^{-1}$.

5. 除(iii)外全是.

6. $\frac{1}{25}\begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}$. (把矩阵看成旋转与膨胀的乘积.)

7. $X=B^{-1}ABC$. 所以当且仅当 A 与 B 交换时, $X=AC$.

8. 这是一个正方形的对称变换群，直角坐标系的原点在这个正方形的中心.

9. 这个群由

$$\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

(逆时针旋转$\frac{2\pi}{3}$)与

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

(关于$O2$-轴的反射)生成.

10. 矩阵 σ_x 与 σ_z 生成了练习 8 中的群. 由 σ_x, σ_y 与 σ_z 生成的群的阶是 16.

11. $(BA)^n=\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

12. 设 A 是非奇异的, 则 $A=A^{-1}(A^2)=A^{-1}(9A)=9I$, 矛盾.

13. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; (iii) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

15. $A=\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 7 & 8 \\ 0 & 1 & -5 & -8 \end{pmatrix}$.

还可能有其它解答, 但是都要含有同一个 2×4 矩阵 R.

16. 利用初等矩阵的乘积是非奇异的这个事实.

18. (i) 逆矩阵是 $-\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$;

(ii) 逆矩阵是 $\frac{1}{10}\begin{pmatrix} 5 & -1 & 5 \\ 0 & 2 & 0 \\ -5 & 3 & -15 \end{pmatrix}$;

(iii) 奇异.

20．逆矩阵是 $\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$.

把矩阵 A 左乘一个置换矩阵相当于置换 A 的行．题中的矩阵对应于置换 (13)(245)。

第十二章

1．(i) 出现并带负号；(ii) 出现并带负号；(iii) 不出现．

2．(i) 0；(ii) -10；(iii) 0；(iv) 0；(v) 10；(vi) -1．

3．$x_1=x_3=k$，$x_2=-2k$，k 是任意数．

4．$|kA|$ 的展开式里的每项都有因子 k^n．

6．(i) -98；(ii) 16；(iii) -4．

7．当且仅当 $\begin{vmatrix} -2-\lambda & 1 \\ 1 & -2-\lambda \end{vmatrix}=0$ 时，方程组有非平凡解．此时 $\lambda=-3$ 或 -1．

8．由 12.2 节定理 VII 的推论，当且仅当对所有 i，$|A_i|\neq 0$ 时，$|A_1\cdots A_k|\neq 0$．

9．$(a-b)(b-c)(c-a)$；也可参考 12.4 节例 16．

10．$|AB|=|I|=1$．由定理 VII 的推论，$|A|\neq 0$，所以 A 是非奇异的，所以 A^{-1} 存在并且 $B=A^{-1}AB=A^{-1}$，B 与 A 可交换．

11．0；两行相等．

13．把 A 的每行乘 -1 可得转置矩阵 A'．因为 A 的行数是奇数，所以 $|A'|=-|A|$．用 12.3 节定理 X．

15．(i) -196；(ii) -294．

16．从第一行，第二行，$\cdots$，第 n 行中依次减去最后一行．这时每行有公因子 $x-a$．而且 $D_{n+1}(x)$ 有因子 $x+na$．证明：依次把第一行，第二行，$\cdots$，第 n 行加到最后一行上，$D_{n+1}(x)=(x-a)^n(x+na)$．

18．当 $x=b,c,d$（两行相同）或当 $x=-b-c-d$（依次把第二行，第三行，第四行加到第一行上）时．

21．$\lambda^2+\frac{3}{2}\lambda-\frac{1}{4}$．

22. 方程的形状是 $ax+by+cz=d$（按第一行展开），$(x,y,z)=(a_i,b_i,c_i)$满足这个方程(两行相等)。

23. $2x-2y-3z=14$.

24. 当且仅当

$$\begin{vmatrix} -2-\lambda & 1 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & 1 & -2-\lambda \end{vmatrix}=0,$$

方程有非平凡解．此时 $\lambda=0$ 或 $\lambda=-3$(用 12.4 节练习 16)．当 $\lambda=0$，$x_1=x_2=x_3=k$(任意)；当 $\lambda=-3$，$x_2=k$，$x_3=l$，$x_1=-k-l$.

25. 当且仅当

$$(a-b)(b-c)(c-a)(a+b+c)=0,$$

方程有非平凡解．当 $a=b, a\neq c$ 时，解是：(i) 若 $a=0$，则 $x=y=0, z=k$(任意)；(ii) 若 $a\neq 0$，则 $y=k$(任意)，$x=(2ac+c^2)k$，$z=-(a+c)k/a$.

26. 当 y_1与 y_2线性相关时，方程 $c_1y_1+c_2y_2=0$ 与 $c_1y_1'+c_2y_2'=0$ 对 c_1 与 c_2一定有非平凡解.

第十三章

1. 反射平面上的非零向量与垂直于这个平面的非零向量分别是对应于特征值 1 与 −1 的特征向量.

2. 关于平面 $x_1-2x_2+2x_3=0$ 的反射(解 $A\mathbf{x}=\mathbf{x}$).

3. 垂直于射影的平面的非零向量是特征值为 0 的特征向量.

4. $(qA)\mathbf{x}=(q\lambda)\mathbf{x}$.

7. 设 $A\mathbf{x}=\lambda\mathbf{x}$，$\mathbf{x}\neq\mathbf{0}$，则 $A^2\mathbf{x}=\lambda^2\mathbf{x}$．所以若 $A^2=A$，则

$$(\lambda^2-\lambda)\mathbf{x}=\mathbf{0}.$$

因为 $\mathbf{x}\neq\mathbf{0}$，得 $\lambda^2-\lambda=0$.

8. AB 的迹$=\sum_{i=1}^{3}\sum_{j=1}^{3}a_{ij}b_{ji}$.

10. 分别是 2 维与 1 维.

11. (i) $\mathscr{S}_{-6}$由 $\begin{pmatrix}-2\\1\end{pmatrix}$ 生成，$\mathscr{S}_{-1}$ 由 $\begin{pmatrix}1\\2\end{pmatrix}$ 生成；

(ii) $\mathscr{S}_1$ 由 $\begin{pmatrix}-2\\1\end{pmatrix}$生成；

(iii) $\mathscr{S}_0$由$\begin{pmatrix}0\\1\\0\end{pmatrix}$生成，$\mathscr{S}_5$由$\begin{pmatrix}-5\\2\\10\end{pmatrix}$生成；

(iv) $\mathscr{S}_0$由$\begin{pmatrix}0\\1\\0\end{pmatrix}$与$\begin{pmatrix}2\\0\\1\end{pmatrix}$生成，$\mathscr{S}_5$由$\begin{pmatrix}1\\0\\-2\end{pmatrix}$生成；

(v) 复特征值 $e^{i\alpha}(=\cos\alpha+i\sin\alpha)$与$e^{-i\alpha}$. 对应的特征向量空间由$\begin{pmatrix}1\\-i\end{pmatrix}$与$\begin{pmatrix}1\\i\end{pmatrix}$生成.

12. 正规方式是(a) $-x_1=2x_2=c\sin(\sqrt{6}t+\alpha)$；

(b) $2x_1=x_2=d\sin(t+\beta)$.

14. $\mathscr{S}_{2+2i}$由$\begin{pmatrix}1\\-1\end{pmatrix}$生成，$\mathscr{S}_2$由$\begin{pmatrix}1\\1\end{pmatrix}$生成，$\bar{A}$的特征值是$2-2i$与$2$.

17. 矩阵(i),(iv)与(v)相似于对角矩阵.

(i) $C^{-1}AC=\begin{pmatrix}-6&0\\0&-1\end{pmatrix}$，这里$C=\begin{pmatrix}-2&1\\1&2\end{pmatrix}$，等等.

18. 例如，

$$C=\begin{pmatrix}1&-2&0\\2&1&-2\\0&2&1\end{pmatrix}.$$

20. $x_1=-c\sin(\sqrt{6}t+\alpha)+\frac{1}{2}d\sin(t+\beta)$,

$x_2=\frac{1}{2}c\sin(\sqrt{6}t+\alpha)+d\sin(t+\beta)$.

21. (i) $C^{-1}AC=A$，这里$C=I$；

(ii) 若$A_1=C_1^{-1}A_2C_1$以及$A_2=C_2^{-1}A_3C_2$，则

$$A_1=(C_2C_1)^{-1}A_3(C_2C_1).$$

22. $(C^{-1}AC)^2=C^{-1}ACC^{-1}AC=C^{-1}A^2C$，等等.

23. D^k是对角阵，ii-元素是λ_i^k.

第十四章

1. 除(ii)外全是.

2. 都是其转置矩阵.

3．$A^{-1}=\frac{1}{9}A$。

4．若 $C^{-1}=C'$，则 $(C')^{-1}=C=(C')'$，由 14.1 节定理 II，C' 是正交矩阵。

5．用练习 4。

6．$\frac{(4x_1+3x_2)^2}{a^2}+\frac{(4x_2-3x_1)^2}{b^2}=25$。

7．轴过原点与点 $P(1,1,0)$，旋转角是 α，$\cos\alpha=\frac{1}{3}$。

8．旋转矩阵相似于行列式值是 1 的矩阵 B（12.4 节例 4）．用 13.3 节定理 VII 的推论．

9．取一个新的直角坐标系 $\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3$，使 $\mathbf{f}_1$ 垂直于反射平面，$\mathbf{f}_2$ 与 $\mathbf{f}_3$ 在反射平面上，C 可以取成：

$$C=\frac{1}{3}\begin{pmatrix}-1 & 2 & 2\\ 2 & 2 & -1\\ -2 & 1 & -2\end{pmatrix},$$

（还有别的解答．）

10．由练习 9 可得．迹 $=1$。

12．两个行列式值是 -1 的矩阵的乘积的行列式值是 1．封闭公理对行列式值是 -1 的矩阵不成立。

13．见 14.2 节练习 10．表示旋转 $\pi/2$ 或 $3\pi/2$ 的矩阵的迹是 1．

14．除(i)外全是．

17．$\mathscr{S}_1$ 由 $\begin{pmatrix}1\\ i\end{pmatrix}$ 生成，$\mathscr{S}_i$ 由 $\begin{pmatrix}1\\ -i\end{pmatrix}$ 生成。

18．见 13.2 节练习 11(v)。

第十五章

1．$T=\frac{1}{2}m\dot{x}_1^2+\frac{1}{4}m\dot{x}_2^2$，$V=k(x_1^2+x_2^2-x_1x_2)$，这里 k 是每段的弹性系数．

2．$T=\frac{1}{2}m(\dot{x}_1^2+2\dot{x}_2^2+\dot{x}_3^2)$，$V=k'(x_1^2+x_2^2+x_3^2-x_1x_2-x_2x_3)$，这里 k' 是每段的弹性系数．（注意，$k'=4k/3$．）

3. $$J=m\begin{pmatrix}8 & -2 & -2\\ -2 & 8 & -2\\ -2 & -2 & 8\end{pmatrix}.$$

4. 动能：(i) $4m\omega^2$，(ii) $2m\omega^2$，(iii) $3m\omega^2$.

角动量(i) $m\omega\begin{pmatrix}-2\\ -2\\ 8\end{pmatrix}$，(ii) $\dfrac{m\omega}{\sqrt{3}}\begin{pmatrix}4\\ 4\\ 4\end{pmatrix}$，(iii) $\dfrac{m\omega}{\sqrt{2}}\begin{pmatrix}6\\ 6\\ 4\end{pmatrix}$.

5. (i) 3 与 -3；(ii) 5 与 20；(iii) -3，0 与 3；

(iv) -1，2 与 5；(v) 2 与 -4.

7. $(\overline{iS})'=(-iS)'=iS$. iS 的特征值是实数．用 13.1 节练习 4.

8. 特征值是 0，$\pm i\sqrt{a^2+b^2+c^2}$.

9. $\mathscr{S}_4$由 $\begin{pmatrix}-5\\ 1\end{pmatrix}$ 生成，$\mathscr{S}_{-2}$由 $\begin{pmatrix}1\\ 1\end{pmatrix}$ 生成.

10. 是；因为它们都是埃尔米特矩阵 iS 的特征向量.

11. $(AB-BA)'=(AB)'-(BA)'=B'A'-A'B'=BA-AB$. 特征值是 $2i$ 与 $-2i$.

13. 数 $y_{r+1},\cdots,y_n$ 是B_1 特征向量的分量，因此它们不全为零.

14. 注意：矩阵(i)与(ii)分别是矩阵 A 与 $A+4I$，因此有相同的特征向量集合，对应特征值为(i)$-3,0,3$ 及(ii)$1,4,7$. 对这二个矩阵，C 可以取成

$$C=\frac{1}{3}\begin{pmatrix}-2 & 2 & 1\\ 2 & 1 & 2\\ 1 & 2 & -2\end{pmatrix},$$

矩阵(iii)，(iv)与(v) 分别是$\frac{1}{3}B$，$\frac{1}{2}(B+3I)$及 $B-I$. 特征值分别是(iii)，$-1,1,1$；(iv) $0,3,3$ 以及(v) $-4,2,2$. 对这几个矩阵，C 可取成

$$C=\begin{pmatrix}1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6}\\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\\ 1/\sqrt{3} & 0 & -2/\sqrt{6}\end{pmatrix}.$$

15. 旋转矩阵R是正交阵．因此，如果它也是对称阵，则 $R^{-1}=R'=R$，所以 $R^2=I$. 所以 $R=I$ 或 R 表示一个半周旋转.

16. 半周旋转的矩阵的迹是-1. 用练习 15.

第十六章

1. $C=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$，这是$C$的一种取法。

2. (i)与(ii)：可取

$$C=\frac{1}{3}\begin{pmatrix}-2 & 2 & 1\\ 2 & 1 & 2\\ 1 & 2 & -2\end{pmatrix}.$$

(iii),(iv)与(v)：可取

$$C=\begin{pmatrix}1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6}\\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\\ 1/\sqrt{3} & 0 & -2/\sqrt{6}\end{pmatrix}.$$

用 15.3 节练习14.)

当 (i) $-x_1=x_2=2x_3=2/3\sqrt{3}$；(iii) $x_1=-x_2=x_3=\frac{1}{3}$；(v) $x_1=-x_2=x_3=1/\sqrt{3}$时，$Q=-1$. 对(ii)与(iv)不可能。

3. $$C=\begin{pmatrix}\frac{1}{2} & 1/2\sqrt{10}\\ \frac{1}{2} & -1/2\sqrt{10}\end{pmatrix}.$$

4. 可取

$$C=\begin{pmatrix}1 & 1/\sqrt{3} & 1/\sqrt{3}\\ 1 & -1/\sqrt{3} & -1/\sqrt{3}\\ 0 & -2/\sqrt{3} & 1/\sqrt{3}\end{pmatrix}.$$

练习 2(i)的二次型可通过变换 $\mathbf{x}=C\mathbf{z}$ 化简成 $z_1^2-z_2^2$，这里

$$C=\frac{1}{3\sqrt{3}}\begin{pmatrix}1 & -2 & 2\sqrt{3}\\ 2 & 2 & \sqrt{3}\\ -2 & 1 & 2\sqrt{3}\end{pmatrix}.$$

5. 主轴的方向是：

$$\begin{pmatrix}1\\1\\1\end{pmatrix},\ \begin{pmatrix}1\\0\\-1\end{pmatrix},\ \begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

对应于这三根主轴，角速度的分量分别是：

(i) $\omega\sqrt{3}, -\omega/\sqrt{2}, \omega/\sqrt{6}$；(ii) $\omega, 0, 0$；(iii) $2\omega/\sqrt{6}, \omega/2, -\omega/2\sqrt{3}$。

动能：(i) $T=\frac{1}{2}m\omega^2(4\cdot(1/\sqrt{3})^2+10(1/\sqrt{2})^2+10(1/\sqrt{6})^2)=4m\omega^2$，等等。

6. $Ch^*=JC\omega^*$，由此得所要的结论。

7. (i) 双曲线，主轴 $x_1+2x_2=0$ 与 $2x_1-x_2=0$；

(ii) 双曲线，中心$\left(-\frac{1}{5}, \frac{3}{5}\right)$，主轴

$$x_1+2x_2=1,\ 2x_1-x_2=-1;$$

(iii) 椭圆，中心(1,1)，主轴 $2x_1+x_2=3$，$x_1-2x_2=-1$；

(iv) 抛物线，顶点$\left(-\frac{1}{5}, -\frac{2}{5}\right)$，主轴 $2x_1+x_2=-1$；

(v) 双曲线，主轴 $x_1-x_2=0$，$x_1+x_2=0$。

9. (i) 正定；(ii) 非负定；(iii) 不定；(iv) 负定。

10. 型(ii)是正定，型(iv)是非负定。

12. (i) $A=\begin{pmatrix}5 & -2\\ -2 & 2\end{pmatrix}^2$；(ii) $A=\begin{pmatrix}9/\sqrt{10} & 3/\sqrt{10}\\ 3/\sqrt{10} & 1/\sqrt{10}\end{pmatrix}^2$。

13. 条件保证特征值的和与乘积是正的，因此二个特征值都是正的。

14. (i) $C=\begin{pmatrix}-1/\sqrt{3} & 1\\ 1/2\sqrt{3} & \frac{1}{2}\end{pmatrix}$，$\mu_1=\frac{2}{3}$，$\mu_2=-2$；

(ii) $C=\begin{pmatrix}2 & 3\\ 1 & 1\end{pmatrix}$，$\mu_1=1$，$\mu_2=4$。

15. 正规坐标是 $y_1=-x_1+3x_2$ 与 $y_2=x_1-2x_2$。

16. (1) 正规方式 $x_2=(\sqrt{3}-1)x_1$，角频率 $\omega=\sqrt{(3-\sqrt{3})q}$ 以及

$x_2=-(\sqrt{3}+1)x_1$，$\omega=\sqrt{(3+\sqrt{3})q}$，这里 $q=k/m$。

(2) 正规方式(i) $x_1=-x_3$，$x_2=0$，$\omega=\sqrt{2q'}$；

(ii) $x_1=x_3=-\frac{1}{2}(1+\sqrt{5})x_2$, $\omega=\sqrt{\frac{1}{2}(3+\sqrt{5})q'}$;

(iii) $x_1=x_3=\frac{1}{2}(\sqrt{5}-1)x_2, \omega=\sqrt{\frac{1}{2}(3-\sqrt{5})q'}$，这里 $q'=k'/m$。

第十七章

1．20，2。

2．5，−5。

3．8，−1。

第十八章

1．

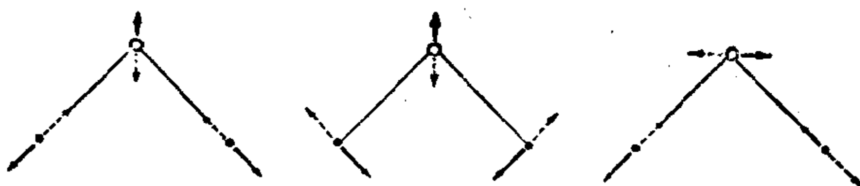


图 103　H_2O 的正规振动方式

3．$\frac{24}{25}c$。

7．(ii) $U'(x',t')=\sin k(x'-ct')$，这里 $k=(1-v/c)/\sqrt{1-v^2/c^2}$。

8．$\frac{\partial^2 U'}{\partial x'^2}=\gamma^2\left(\frac{\partial^2 U}{\partial x^2}-2\frac{v}{c^2}\frac{\partial^2 U}{\partial x\partial t}+\frac{v^2}{c^4}\frac{\partial^2 U}{\partial t^2}\right)$ 以及

$$\frac{\partial^2 U'}{\partial t'^2}=\gamma^2\left(v^2\frac{\partial^2 U}{\partial x^2}-2v\frac{\partial^2 U}{\partial x\partial t}+\frac{\partial^2 U^2}{\partial t^2}\right)，等等。$$

10．各分量都是 1 的列向量是这样的特征向量。

11．(i) $\frac{3}{16}$；(ii) $p_{31}=p_{32}=0$；(iii) k 步后不吸收到状态 3 的概率是 $\left(\frac{3}{4}\right)^k$。

12

$$P^k=\frac{1}{4}\begin{pmatrix}3\alpha+\beta & \alpha-\beta & 4-4\alpha\\ 3\alpha-3\beta & \alpha+3\beta & 4-4\alpha\\ 0 & 0 & 4\end{pmatrix}，\quad 这里\ \alpha=\left(\frac{3}{4}\right)^k，$$

$\beta=\left(-\frac{1}{4}\right)^{n}$.

17．用练习 13 至 16。

19． $\omega>\frac{1}{2}\sqrt{LC}$.

20．若 a 是实数，则 $\sin\beta=0$，于是 $a=\pm\cosh\alpha$。

附录

[s=对称关系，t=传递关系，r=自反关系，e=等价关系.]

1．(i) s；(ii) t；(iii) t,r.

2．s.

3．e. 每类是 $a+bx+(x^2+1)q(x)$ 这种多项式的全体，$q(x)$是任意的。

4．e. 每类是陪集$\mathscr{H}a$=所有的 ha，h 在$\mathscr{H}$中.

5．(i) t,r；(ii) e. 每类是具有相同秩的矩阵的集合。（矩阵的秩是它的行生成的向量空间的维数.）

汉英对照名词索引

（按笔划排列，笔划相同的字按起笔类型以丶 一丨 ㇏顺序排列）

一至三划

四　划

五　划

六　划

七 划

九 划

十 划

十 一 划

十 二 划

十 三 划

十四划以上